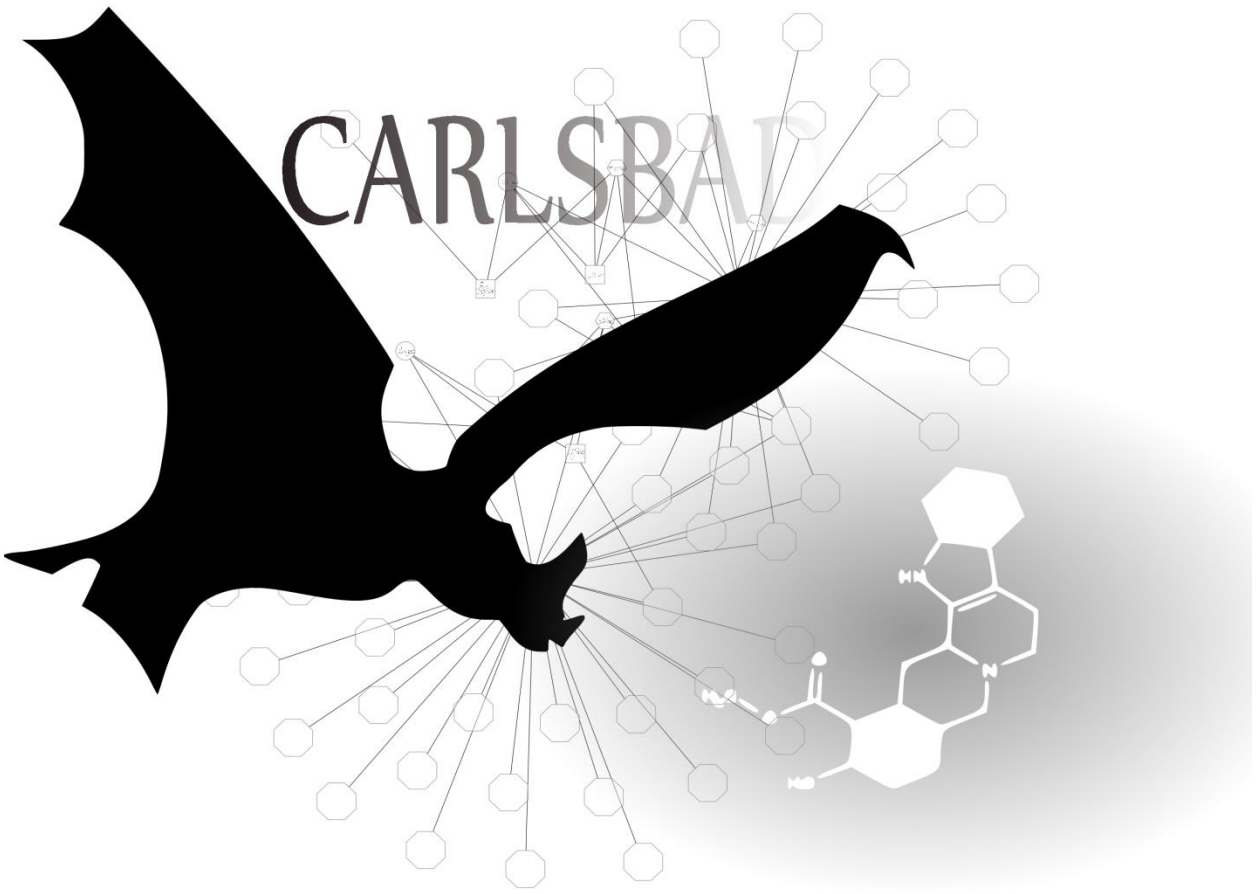
# CARLSBAD User Manual & Tutorial




**The Power to Explore Biological Networks via Chemical Patterns**

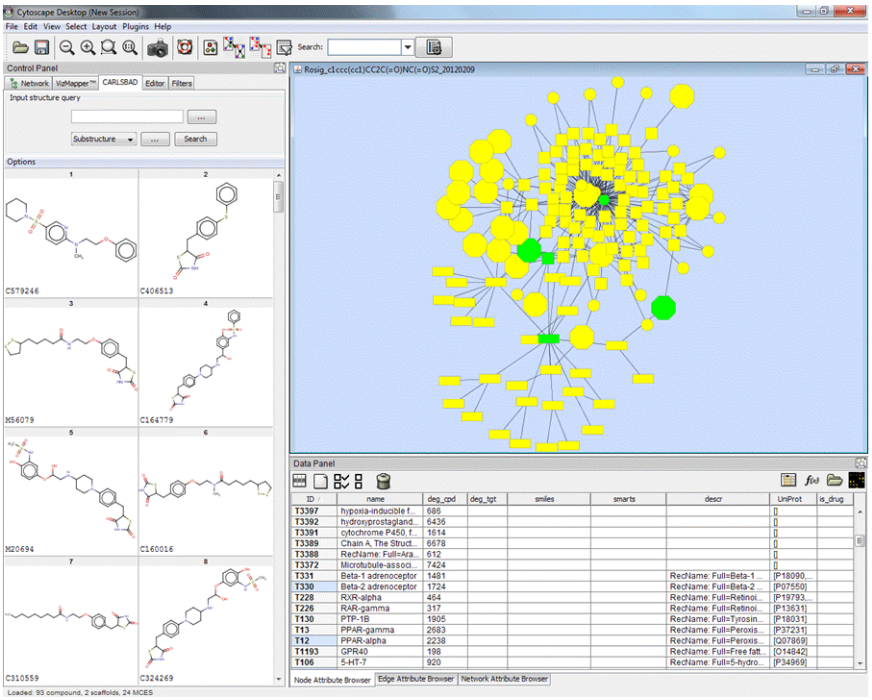
**carlsbad.health.unm.edu**

Copyright (C) 2012 Jarrett Hines-Kay

Permission is granted to copy, distribute and/or modify this document under the terms of the GNU Free Documentation License, Version 1.3 or later, published by the Free Software Foundation; with no Invariant Sections, with the Front-Cover Text being CARLSBAD User Manual, and no Back-Cover Texts. A copy of the license is included in the section entitled, "GNU Free Documentation License," of this manual.

This is a reference manual for the CARLSBAD platform which consists of a database, a database query tool called SNAKE and a plugin to Cytoscape for network visualization and analysis. The manual is constructed around a tutorial to help users better understand CARLSBADs capabilities and how the system can be utilized for hypothesis generation. The CARLSBAD platform is free for academic and non-commercial usage however you must apply for a User License (**carlsbad.health.unm.edu/cgi-bin/cb_user_app.cgi**) and this can take several days. In order for you to test the platform we have made the two files that the tutorial uses available on the CARLSBAD project homepage (**carlsbad.health.unm.edu**). These files can be obtained by clicking the "CARLSBAD Community" link and downloading the two CARLSBAD tutorial XGMML files under the "Downloads" heading. In addition, there is a glossary of definitions that can be found at the end of the manual and each item that is defined appears in red font the first time it is mentioned in the manual content.

CARLSBAD is a database and knowledge inference system that integrates multiple bioactivity datasets in order to provide researchers with novel capabilities for the mining and exploration of available structure activity relationships (SAR) throughout chemical biology space. The powerful cheminformatics and pattern recognition algorithms combined with network analysis methodology allow researchers using CARLSBAD to generate hypotheses involving the relationships between biological targets, chemical compounds and their common chemical patterns (CCPs). The CARLBAD platform consists of a web-based extraction tool called SNAKE and a network analysis plugin for Cytoscape called CARLSBAD. In addition, the use of Cytoscape enables researchers to infer novel associations via the powerful biological network visualization environment that Cytoscape provides.




CARLSBAD is offered **FREE** for academic and non-commercial research. In order to **APPLY** for CARLSBAD user group membership (**carlsbad.health.unm.edu/cgi-bin/cb_user_app.cgi**), obtain system access, access supporting documentation and downloads, as well as learn more about the CARLSBAD platform and UNM's Translational Informatics Division, please visit the CARLSBAD website at **carlsbad.health.unm.edu.**

Questions can be answered by emailing CARLSBAD at **info-carlsbad@poblano.health.unm.edu.**

CARLSBAD was developed at the University of New Mexico School of Medicine Translational Informatics Division by Tudor Oprea, Cristian Bologa, Stephen Mathias, Oleg Ursu, Jeremy Yang, Gergely Zahoranszky-Kohalmi, Jerome Abear and Jarrett Hines-Kay and funded by NIH grant GM095952.

**TABLE OF CONTENTS**


1. **Installation and Downloading Procedures** **5**
   - Obtaining a CARLSBAD user group license
   - Installing Cytoscape
   - Obtaining and Installing a ChemAxon license
   - Loading the CARLSBAD plugin to Cytoscape

2. **SNAKE: SubNet Application Kit and Extractor for CARLSBAD** **7**
   - Getting familiar with SNAKE
   - Query Options
   - Result Options
   - Cytoscape Web
   - Additional Options

3. **The CARLSBAD Plugin for Cytoscape** **13**
   - Importing a file downloaded from SNAKE into the CARLSBAD plugin for Cytoscape
   - CARLSBAD Node Attributes and Cytoscape Viewing Panels
   - CARLSBAD Edge Attributes
   - Using the CARLSBAD plugin options: Filtering tools
     i. Remove unmergeable CCPs
     ii. Remove promiscuous compounds
     iii. Remove specified activity edges
   - Using the CARLSBAD plugin options: Additional tools
     i. Apply force-directed similarity weighted layout
     ii. Select structural class
     iii. Connect structural patterns to targets

4. **Merging CARLSBAD Networks** **28**

5. **Next Steps** **31**
   - Literature Searches
   - Additional Biocomputing Tools for Filtering
     i. BadApple
     ii. SmartsFilter

6. **Validation** **37**

7. **Appendix** **38**
   - Applying Different Visual Styles to a CARLSBAD Network

8. **Glossary** **41**

9. **CARLSBAD References** **44**

10. **GNU Free Documentation License** **45**

1. **Installation and Downloading Procedures** - To begin using the CARLSBAD platform, you will need to obtain a CARLSBAD group license (free for academic and non-commercial usage), install **Cytoscape** (free), a ChemAxon license (free for some categories such as academics and non-commercial usage) and load the **CARLSBAD plugin** to Cytoscape. This section will focus on how to download and install these applications.

    - **Obtaining a CARLSBAD User Group License** - In order to begin using the **subnet** extractor tool, **SNAKE** (see section 2 of this manual), please visit **carlsbad.health.unm.edu/cgi-bin/cb_user_app.cgi** to apply for a license which can take several days to approve.

    - **Installing Cytoscape** - Cytoscape is an open source bioinformatics software platform that CARLSBAD is able to utilize, via the plugin, for powerful network visualization. To install Cytoscape, go to the Cytoscape website at **www.cytoscape.org/download.html** and follow the download instructions (version 2.8.2 or above).

    - **Obtaining a ChemAxon license** - In order to obtain a ChemAxon license you must: i) decide which type of license to request; ii) request the license; and, iii) follow standard license installation procedures. This information can be found at **www.chemaxon.com/my-chemaxon/my-academic-license/.**

    - **Installing the ChemAxon license** - Once you have obtained a ChemAxon license, copy the retrieved file, "license.cxl," to one of the following directories according to the operating system on your machine:

        i. **Windows 7**
            - C:\Users*your_username*\chemaxon\

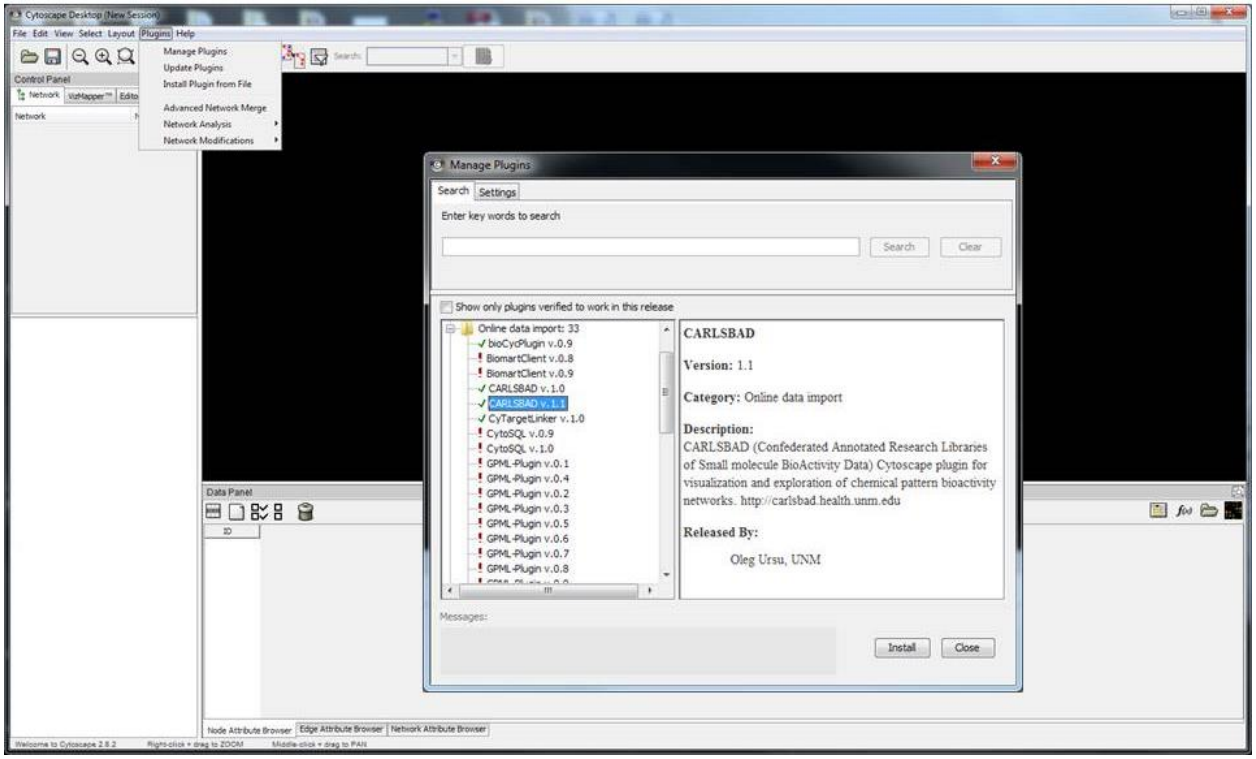
        ii. **UNIX**
            - /home/*your_username*/.chemaxon/

        iii. **Apple**
            - *$HOME/your_username*/.chemaxon/
                - Substitute *$HOME* with the path to your home folder.

        *(Note: You may have to create the directory yourself).*

    - **Installing a Java Virtual Machine** - In addition, we recommend that you ensure your computer is equipped as a Java Virtual Machine (SE 6 Update 33 JRE or recent series 6 JRE is recommended). Java updates can be found at, **www.oracle.com/technetwork/java/javase/downloads/jre6-downloads-1637595.html**.

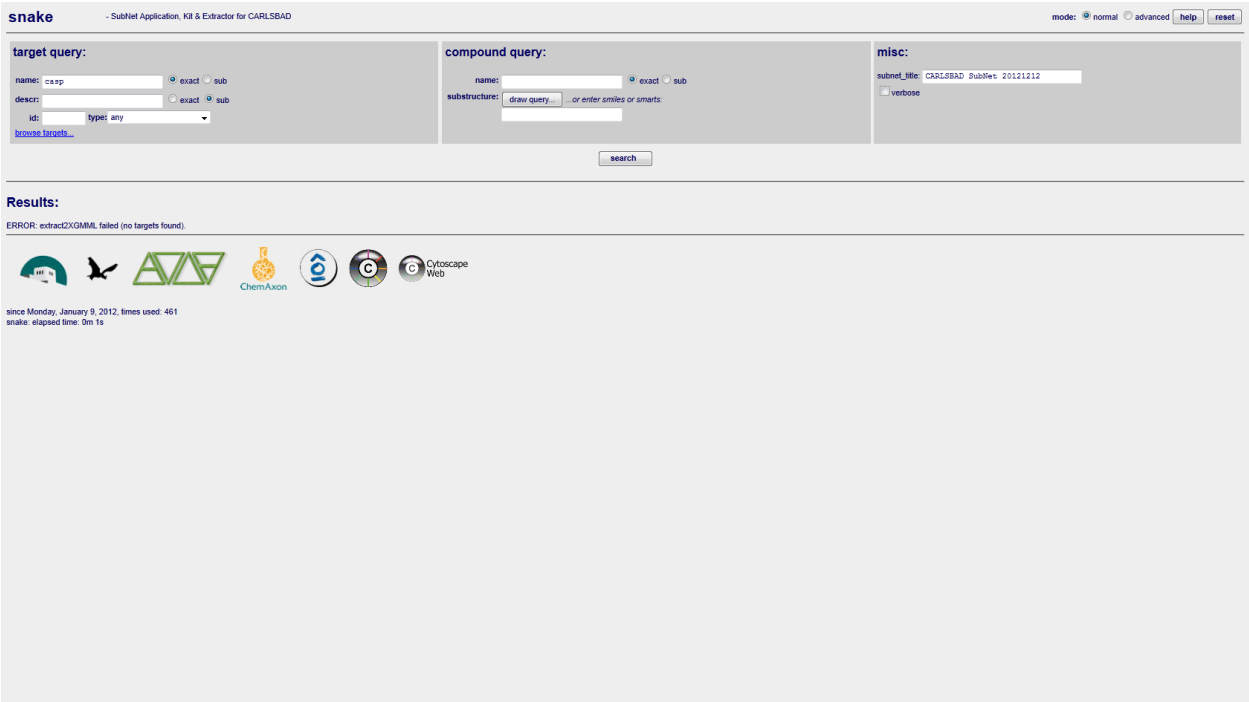
- **Loading the CARLSBAD plugin to Cytoscape** - After installing Cytoscape, you will need to install the CARLSBAD plugin by launching Cytoscape, going to the "Plugins" menu in Cytoscape and selecting "Manage Plugins" as depicted in the screenshot below. This will launch the "Manage Plugins" pop-up window where you can expand the "Online data import:" subtree by clicking on the "+" next the folder icon, select the "CARLSBAD v.1.1" plugin (highlighted below) and click the "Install" button.




The installation process will take several minutes. Once installed, click the "Close" button and you are now ready to begin using the CARLSBAD plugin to Cytoscape, a process that will begin with a subnet query in SNAKE, which is described in the next section.

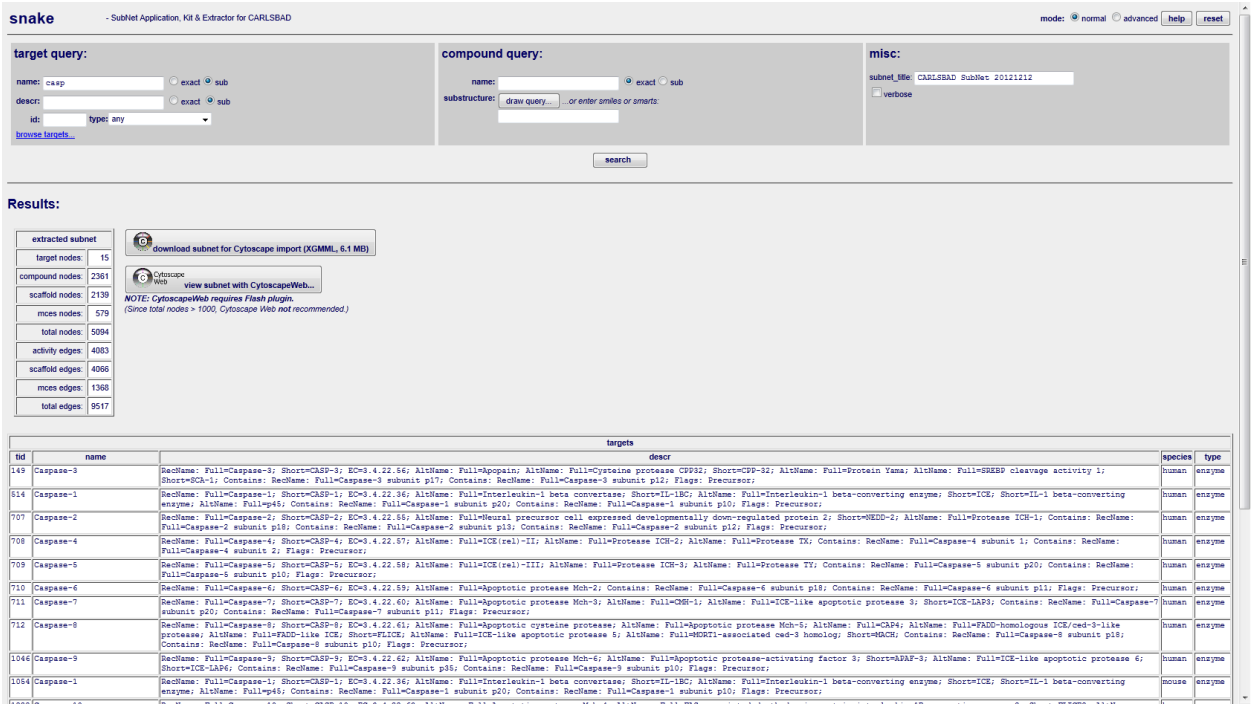
2. **SNAKE: SubNet Application Kit and Extractor for CARLSBAD -** SNAKE is a web-based tool where you can perform **target** or **compound** based queries. After performing your query, SNAKE delivers target results for your **network**, an option to view your network with **Cytoscape Web** and the ability to download a network file that can be imported into Cytoscape for advanced network visualization and analysis with the CARLSBAD plugin. To begin using SNAKE, please apply for a CARLSBAD user group license offered free for academic and non-commercial usage at **carlsbad.health.unm.edu/cgi-bin/cb_user_app.cgi.**

    - **Getting familiar with SNAKE -** SNAKE can be accessed through the CARLSBAD homepage at **carlsbad.health.unm.edu** or via **carlsbad.health.unm.edu/tomcat/biocomp/snake**. Once the web page is loaded, you will notice the “target query” options in the left box and the “compound query” options in the center box. To perform a substructure or **common chemical pattern (CCP)** query, you can either draw the structure with the button in the “compound query” box or enter the **SMILES/SMARTS** in the space below the "draw query" button. The final box, “misc,” allows you to name the network file that you will download from SNAKE and open in Cytoscape for analysis with the CARLSBAD plugin (described further in the “CARLSBAD Plugin to Cytoscape” section of this manual).




*(Note: There is a glossary at the end of this manual with many definitions for items mentioned in the manual and these defined items appear in red font the first time they are mentioned in the manual content; i.e., the "common chemical patterns (CCP)" example found above).*

- **Query Options** - In the “target query” box, you can search by the target name (“name”), a description of the target (“descr”) or one of the 19 **identifiers** (“id”), which can be found in the drop-down menu for “type.” In the “compound query” box, you can search by name (“name”), a drawn structure (“structure”) or SMILES/SMARTS. In addition, you will notice the options “exact” and “sub” after entering text. These allow you to narrow or broaden your search. For example, you can see in the previous screenshot that a search for "casp" (short for caspase) retrieves no results when “exact” is selected, however ….


snake
- SubNet Application, Kit & Extractor for CARLSBAD
mode: normal advanced help reset
target query:
name: casp exact sub
descr: exact sub
id: type: any
browse targets...
compound query:
name: exact sub
substructure: draw query... ...or enter smiles or smarts:
misc:
subnet_title: CARLSBAD SubNet 20121212
verbose
search
Results:
extracted subnet
target nodes: 15
compound nodes: 2361
scaffold nodes: 2139
mces nodes: 579
total nodes: 5094
activity edges: 4083
scaffold edges: 4066
mces edges: 1368
total edges: 9517
download subnet for Cytoscape import (XGMML, 6.1 MB)
view subnet with CytoscapeWeb...
NOTE: CytoscapeWeb requires Flash plugin.
(Since total nodes > 1000, Cytoscape Web not recommended.)
targets
tid name descr species type
149 Caspase-3 RecName: Full=Caspase-3; Short=CASP-3; EC=3.4.22.56; AltName: Full=Apopain; AltName: Full=Cysteine protease CPP32; Short=CPP-32; AltName: Full=Protein Yama; AltName: Full=SREBP cleavage activity 1; Short=SCA-1; Contains: RecName: Full=Caspase-3 subunit p17; Contains: RecName: Full=Caspase-3 subunit p12; Flags: Precursor; human enzyme
514 Caspase-1 RecName: Full=Caspase-1; Short=CASP-1; EC=3.4.22.36; AltName: Full=Interleukin-1 beta convertase; Short=IL-1BC; AltName: Full=Interleukin-1 beta-converting enzyme; Short=ICE; Short=IL-1 beta-converting enzyme; AltName: Full=p45; Contains: RecName: Full=Caspase-1 subunit p20; Contains: RecName: Full=Caspase-1 subunit p10; Flags: Precursor; human enzyme
707 Caspase-2 RecName: Full=Caspase-2; Short=CASP-2; EC=3.4.22.55; AltName: Full=Neural precursor cell expressed developmentally down-regulated protein 2; Short=NEDD-2; AltName: Full=Protease ICH-1; Contains: RecName: Full=Caspase-2 subunit p18; Contains: RecName: Full=Caspase-2 subunit p13; Contains: RecName: Full=Caspase-2 subunit p12; Flags: Precursor; human enzyme
708 Caspase-4 RecName: Full=Caspase-4; Short=CASP-4; EC=3.4.22.57; AltName: Full=ICE(rel)-II; AltName: Full=Protease ICH-2; AltName: Full=Protease TX; Contains: RecName: Full=Caspase-4 subunit 1; Contains: RecName: Full=Caspase-4 subunit 2; Flags: Precursor; human enzyme
709 Caspase-5 RecName: Full=Caspase-5; Short=CASP-5; EC=3.4.22.58; AltName: Full=ICE(rel)-III; AltName: Full=Protease ICH-3; AltName: Full=Protease TY; Contains: RecName: Full=Caspase-5 subunit p20; Contains: RecName: Full=Caspase-5 subunit p10; Flags: Precursor; human enzyme
710 Caspase-6 RecName: Full=Caspase-6; Short=CASP-6; EC=3.4.22.59; AltName: Full=Apoptotic protease Mch-2; Contains: RecName: Full=Caspase-6 subunit p18; Contains: RecName: Full=Caspase-6 subunit p11; Flags: Precursor; human enzyme
711 Caspase-7 RecName: Full=Caspase-7; Short=CASP-7; EC=3.4.22.60; AltName: Full=Apoptotic protease Mch-3; AltName: Full=CMH-1; AltName: Full=ICE-like apoptotic protease 3; Short=ICE-LAP3; Contains: RecName: Full=Caspase-7 subunit p20; Contains: RecName: Full=Caspase-7 subunit p11; Flags: Precursor; human enzyme
712 Caspase-8 RecName: Full=Caspase-8; Short=CASP-8; EC=3.4.22.61; AltName: Full=Apoptotic cysteine protease; AltName: Full=Apoptotic protease Mch-5; AltName: Full=CAP4; AltName: Full=FADD-homologous ICE/ced-3-like protease; AltName: Full=FADD-like ICE; Short=FLICE; AltName: Full=ICE-like apoptotic protease 5; AltName: Full=MORT1-associated ced-3 homolog; Short=MACH; Contains: RecName: Full=Caspase-8 subunit p18; Contains: RecName: Full=Caspase-8 subunit p10; Flags: Precursor; human enzyme
1046 Caspase-9 RecName: Full=Caspase-9; Short=CASP-9; EC=3.4.22.62; AltName: Full=Apoptotic protease Mch-6; AltName: Full=Apoptotic protease-activating factor 3; Short=APAF-3; AltName: Full=ICE-like apoptotic protease 6; Short=ICE-LAP6; Contains: RecName: Full=Caspase-9 subunit p35; Contains: RecName: Full=Caspase-9 subunit p10; Flags: Precursor; human enzyme
1054 Caspase-1 RecName: Full=Caspase-1; Short=CASP-1; EC=3.4.22.36; AltName: Full=Interleukin-1 beta convertase; Short=IL-1BC; AltName: Full=Interleukin-1 beta-converting enzyme; Short=ICE; Short=IL-1 beta-converting enzyme; AltName: Full=p45; Contains: RecName: Full=Caspase-1 subunit p20; Contains: RecName: Full=Caspase-1 subunit p10; Flags: Precursor; mouse enzyme

… By changing the option from “exact” to “sub” will allow you to broaden your search and, as you can see from the “extracted subnet” in the “Results” above, the same “casp” query performed as “sub” retrieved 15 targets.

*(Note: In SNAKE's "target query" box there is a link towards the bottom that can take you to a searchable html listing of all the targets contained in CARLSBAD).*

- **Result Options** - If you scroll down to the “Results” section, you will find some basic **attributes** in the “extracted subnet” section, information on the targets in the “targets” section and three buttons. The button at the bottom of the page, “download targets,” downloads a ".csv" file (which can be opened with a program like Excel by Microsoft) that contains the target information depicted in this screenshot as well as additional “identifier” information. The button at the top, “download subnet for Cytoscape import,” allows you to download a SNAKE file that can be opened in Cytoscape for visualization and analysis with the CARLSBAD plugin (described further in the “CARLSBAD Plugin to Cytoscape” section of this manual). Finally, the button titled, “view subnet with CytoscapeWeb,” allows you to quickly visualize your network without having to import it into Cytoscape.

**Results:**

| extracted subnet | |
|---|---|
| target nodes: | 15 |
| compound nodes: | 2361 |
| scaffold nodes: | 2139 |
| mces nodes: | 579 |
| total nodes: | 5094 |
| activity edges: | 4083 |
| scaffold edges: | 4066 |
| mces edges: | 1368 |
| total edges: | 9517 |

download subnet for Cytoscape import (XGMML, 6.1 MB)

view subnet with CytoscapeWeb...

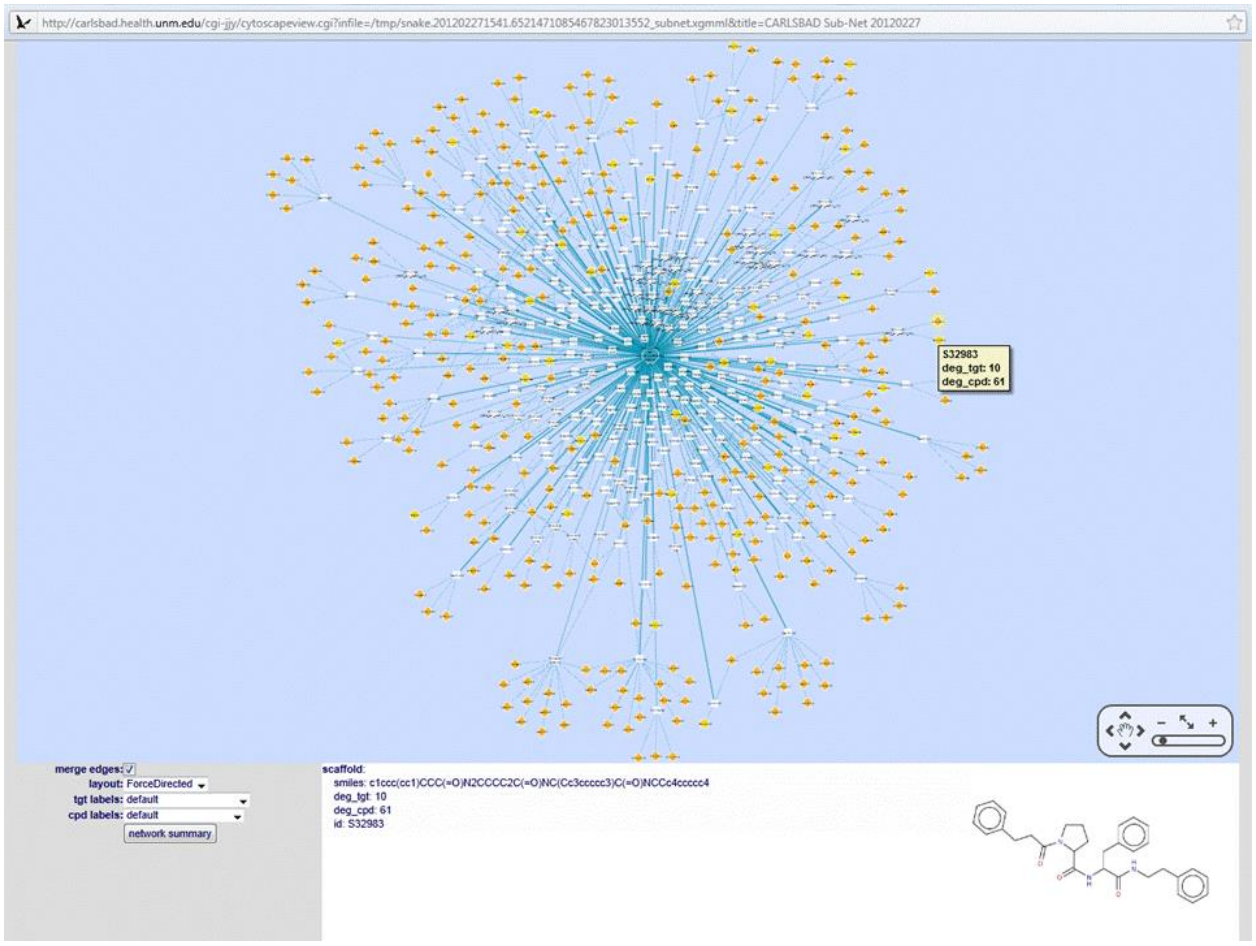
*NOTE: CytoscapeWeb requires Flash plugin.*
*(Since total nodes > 1000, Cytoscape Web not recommended.)*

targets

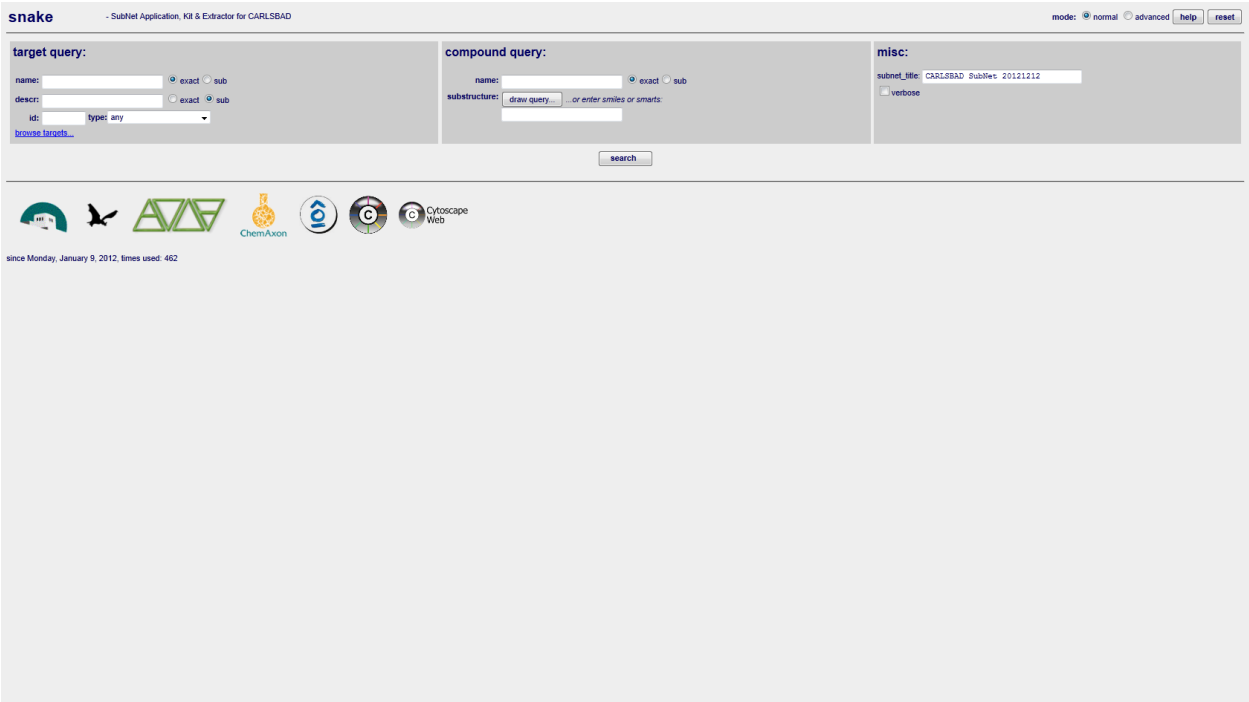
| tid | name | descr | species | type |
|---|---|---|---|---|
| 149 | Caspase-3 | RecName: Full=Caspase-3; Short=CASP-3; EC=3.4.22.56; AltName: Full=Apopain; AltName: Full=Cysteine protease CPP32; Short=CPP-32; AltName: Full=Protein Yama; AltName: Full=SREBP cleavage activity 1; Short=SCA-1; Contains: RecName: Full=Caspase-3 subunit p17; Contains: RecName: Full=Caspase-3 subunit p12; Flags: Precursor; | human | enzyme |
| 514 | Caspase-1 | RecName: Full=Caspase-1; Short=CASP-1; EC=3.4.22.36; AltName: Full=Interleukin-1 beta convertase; Short=IL-1BC; AltName: Full=Interleukin-1 beta-converting enzyme; Short=ICE; Short=IL-1 beta-converting enzyme; AltName: Full=p45; Contains: RecName: Full=Caspase-1 subunit p20; Contains: RecName: Full=Caspase-1 subunit p10; Flags: Precursor; | human | enzyme |
| 707 | Caspase-2 | RecName: Full=Caspase-2; Short=CASP-2; EC=3.4.22.55; AltName: Full=Neural precursor cell expressed developmentally down-regulated protein 2; Short=NEDD-2; AltName: Full=Protease ICH-1; Contains: RecName: Full=Caspase-2 subunit p18; Contains: RecName: Full=Caspase-2 subunit p13; Contains: RecName: Full=Caspase-2 subunit p12; Flags: Precursor; | human | enzyme |
| 708 | Caspase-4 | RecName: Full=Caspase-4; Short=CASP-4; EC=3.4.22.57; AltName: Full=ICE(rel)-II; AltName: Full=Protease ICH-2; AltName: Full=Protease TX; Contains: RecName: Full=Caspase-4 subunit 1; Contains: RecName: Full=Caspase-4 subunit 2; Flags: Precursor; | human | enzyme |
| 709 | Caspase-5 | RecName: Full=Caspase-5; Short=CASP-5; EC=3.4.22.58; AltName: Full=ICE(rel)-III; AltName: Full=Protease ICH-3; AltName: Full=Protease TY; Contains: RecName: Full=Caspase-5 subunit p20; Contains: RecName: Full=Caspase-5 subunit p10; Flags: Precursor; | human | enzyme |
| 710 | Caspase-6 | RecName: Full=Caspase-6; Short=CASP-6; EC=3.4.22.59; AltName: Full=Apoptotic protease Mch-2; Contains: RecName: Full=Caspase-6 subunit p18; Contains: RecName: Full=Caspase-6 subunit p11; Flags: Precursor; | human | enzyme |
| 711 | Caspase-7 | RecName: Full=Caspase-7; Short=CASP-7; EC=3.4.22.60; AltName: Full=Apoptotic protease Mch-3; AltName: Full=CMH-1; AltName: Full=ICE-like apoptotic protease 3; Short=ICE-LAP3; Contains: RecName: Full=Caspase-7 subunit p20; Contains: RecName: Full=Caspase-7 subunit p11; Flags: Precursor; | human | enzyme |
| 712 | Caspase-8 | RecName: Full=Caspase-8; Short=CASP-8; EC=3.4.22.61; AltName: Full=Apoptotic cysteine protease; AltName: Full=Apoptotic protease Mch-5; AltName: Full=CAP4; AltName: Full=FADD-homologous ICE/ced-3-like protease; AltName: Full=FADD-like ICE; Short=FLICE; AltName: Full=ICE-like apoptotic protease 5; AltName: Full=MORT1-associated ced-3 homolog; Short=MACH; Contains: RecName: Full=Caspase-8 subunit p18; Contains: RecName: Full=Caspase-8 subunit p10; Flags: Precursor; | human | enzyme |
| 1046 | Caspase-9 | RecName: Full=Caspase-9; Short=CASP-9; EC=3.4.22.62; AltName: Full=Apoptotic protease Mch-6; AltName: Full=Apoptotic protease-activating factor 3; Short=APAF-3; AltName: Full=ICE-like apoptotic protease 6; Short=ICE-LAP6; Contains: RecName: Full=Caspase-9 subunit p35; Contains: RecName: Full=Caspase-9 subunit p10; Flags: Precursor; | human | enzyme |
| 1054 | Caspase-1 | RecName: Full=Caspase-1; Short=CASP-1; EC=3.4.22.36; AltName: Full=Interleukin-1 beta convertase; Short=IL-1BC; AltName: Full=Interleukin-1 beta-converting enzyme; Short=ICE; Short=IL-1 beta-converting enzyme; AltName: Full=p46; Contains: RecName: Full=Caspase-1 subunit p20; Contains: RecName: Full=Caspase-1 subunit p10; Flags: Precursor; | mouse | enzyme |
| 1222 | Caspase-10 | RecName: Full=Caspase-10; Short=CASP-10; EC=3.4.22.63; AltName: Full=Apoptotic protease Mch-4; AltName: Full=FAS-associated death domain protein interleukin-1B-converting enzyme 2; Short=FLICE2; AltName: Full=ICE-like apoptotic protease 4; Contains: RecName: Full=Caspase-10 subunit p23/17; Contains: RecName: Full=Caspase-10 subunit p12; Flags: Precursor; | human | enzyme |
| 3443 | Caspase-4 | RecName: Full=Caspase-4; Short=CASP-4; EC=3.4.22.64; AltName: Full=Caspase-11; Short=CASP-11; AltName: Full=Protease ICH-3; Contains: RecName: Full=Caspase-4 subunit p10; Contains: RecName: Full=Caspase-4 subunit p20; Flags: Precursor; | mouse | enzyme |
| 3444 | Caspase-14 | RecName: Full=Caspase-14; Short=CASP-14; EC=3.4.22.-; Contains: RecName: Full=Caspase-14 subunit p19; Contains: RecName: Full=Caspase-14 subunit p10; Flags: Precursor; | human | enzyme |
| 3784 | caspase-3 | null | rat | protein |
| 3787 | caspase 7, apoptosis-related cysteine peptidase | null | human | protein |

download targets (CSV, 300.3 KB)

- **Cytoscape Web -** Cytoscape Web offers network visualization as well as additional attribute information and structural depiction, which can be obtained by selecting a **node**. Below is a network depiction generated using the “view subnet with CytoscapeWeb” button found in the "Results" section of SNAKE after performing a query (see previous screenshot).



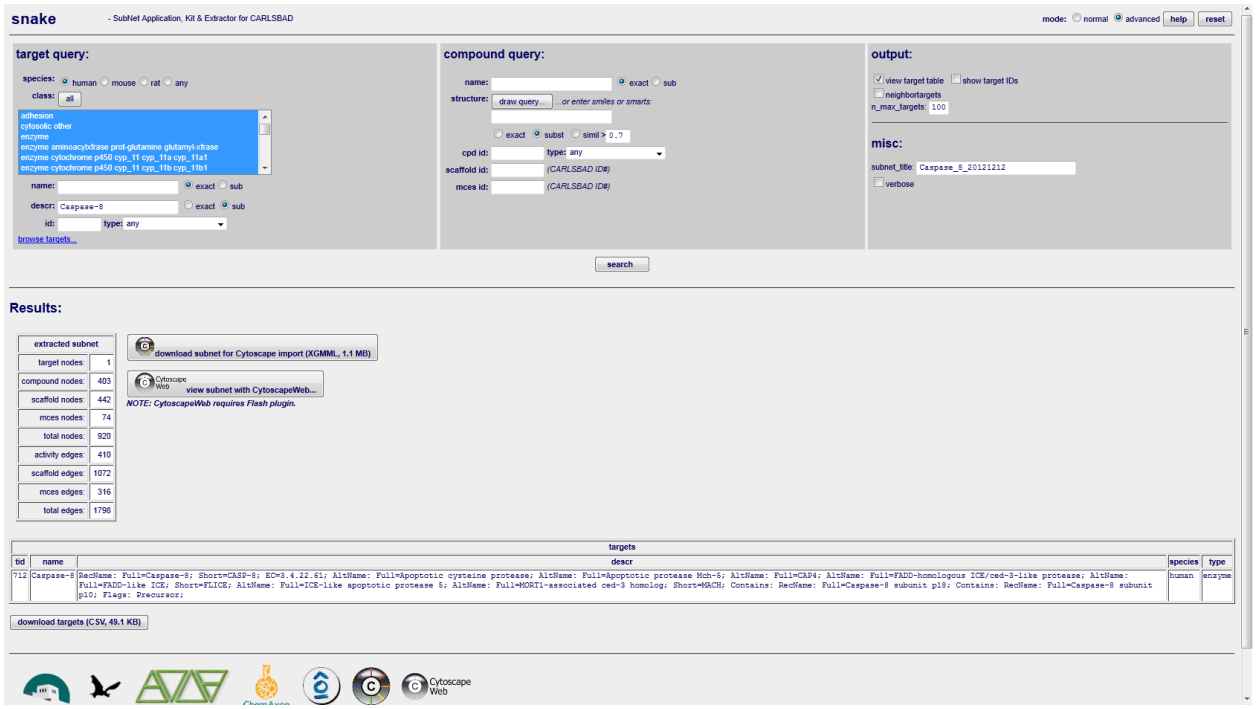
- **Additional Options -** SNAKE provides a few additional options that can be found in the top right region of the SNAKE window. The “reset” button clears all fields and can be useful when analyzing multiple networks, there is a “help” button for assistance and there are selections for “mode” query type ...




... Thus far, the screenshots of SNAKE have been for the “normal” query mode. Below you can see some additional options available when selecting the “advanced” query mode in the top right part of the window ...

*(Note: When using SNAKE to generate a file for import into Cytoscape, naming your file under the "misc:" heading in SNAKE is highly recommended as it is easy to get confused as to which files you are working with despite the date stamp that is automatically applied to the name in SNAKE. In addition, you must name your file before clicking the “search” button in order for your filename to appear in the file you download).*

... For target queries, the expert mode allows you to filter your search by species (human, mouse and rat are default) and target class. In the "compound query box," the expert mode allows for the ability to search compounds and CCPs by CARLSBAD ID. *This feature can be useful, but you should be aware that CARLSBAD IDs are not static and can change, so identification by SMARTS or SMILES is recommended...*

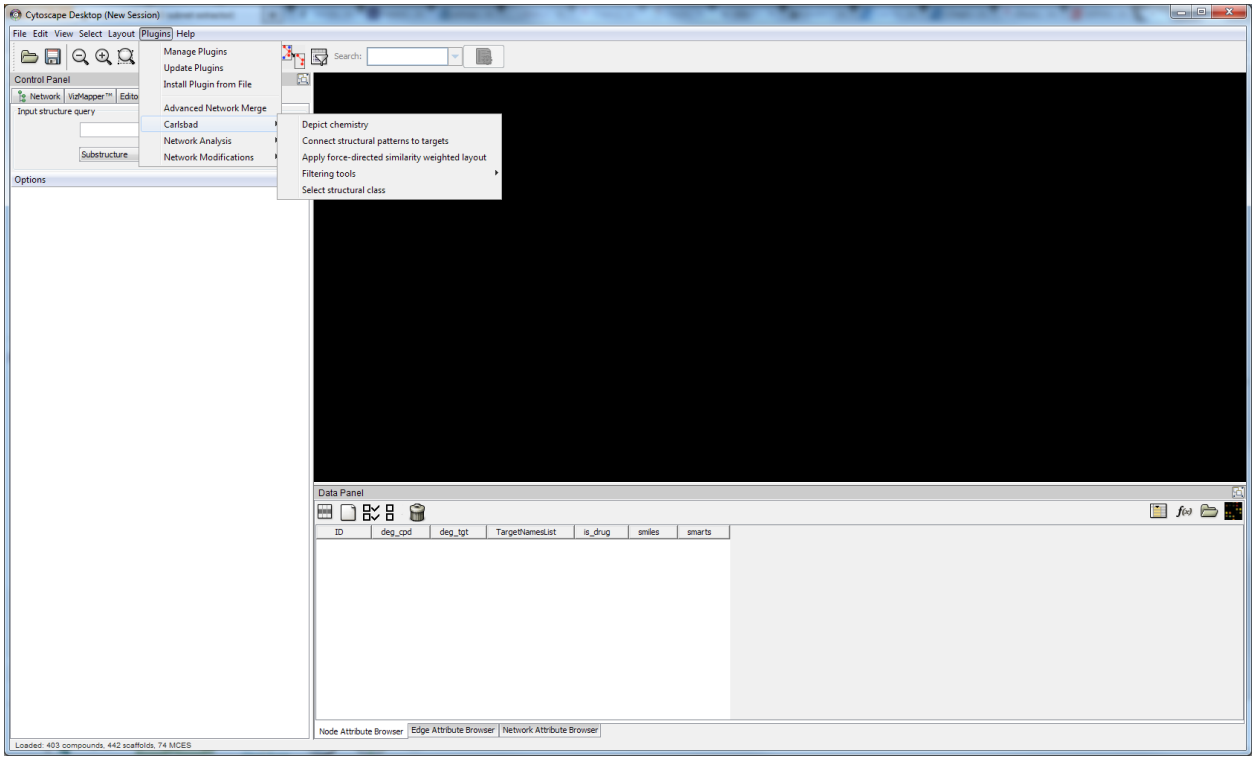

snake
- SubNet Application, Kit & Extractor for CARLSBAD
mode: normal advanced help reset
target query:
species: human mouse rat any
class: all
adhesion
cytosolic other
enzyme
enzyme aminoacyltrase prot-glutamine glutamyl-xfrase
enzyme cytochrome p450 cyp_11 cyp_11a cyp_11a1
enzyme cytochrome p450 cyp_11 cyp_11b cyp_11b1
name: exact sub
descr: Caspase-8 exact sub
id: type: any
browse targets...
compound query:
name: exact sub
structure: draw query... ...or enter smiles or smarts:
exact subst simil > 0.7
cpd id: type: any
scaffold id: (CARLSBAD ID#)
mces id: (CARLSBAD ID#)
output:
view target table show target IDs
neighbortargets
n_max_targets: 100
misc:
subnet_title: Caspase_8_20121212
verbose
search
Results:
extracted subnet
target nodes: 1
compound nodes: 403
scaffold nodes: 442
mces nodes: 74
total nodes: 920
activity edges: 410
scaffold edges: 1072
mces edges: 316
total edges: 1798
download subnet for Cytoscape import (XGMML, 1.1 MB)
view subnet with CytoscapeWeb...
NOTE: CytoscapeWeb requires Flash plugin.
targets
tid name descr species type
712 Caspase-8 RecName: Full=Caspase-8; Short=CASP-8; EC=3.4.22.61; AltName: Full=Apoptotic cysteine protease; AltName: Full=Apoptotic protease Mch-5; AltName: Full=CAP4; AltName: Full=FADD-homologous ICE/ced-3-like protease; AltName: Full=FADD-like ICE; Short=FLICE; AltName: Full=ICE-like apoptotic protease 5; AltName: Full=MORT1-associated ced-3 homolog; Short=MACH; Contains: RecName: Full=Caspase-8 subunit p18; Contains: RecName: Full=Caspase-8 subunit p10; Flags: Precursor; human enzyme
download targets (CSV, 49.1 KB)
Cytoscape Web

... Finally, there are some additional "output" options above the "misc" box such as "n_max_targets," which allow you to alter the amount of results returned in case of query truncation (a warning that will appear if your search yields more than the default 100 targets).

*(Note: This manual will describe CARLSBAD features in the context of a tutorial example that requires you to download the human Caspase-8 target file from SNAKE, depicted above, and import it into Cytoscape for analysis with the CARLSBAD plugin, described in the preceding section. Since CARLSBAD is an ongoing project, it is possible that some of your results will vary slightly from the screenshots in this manual).*

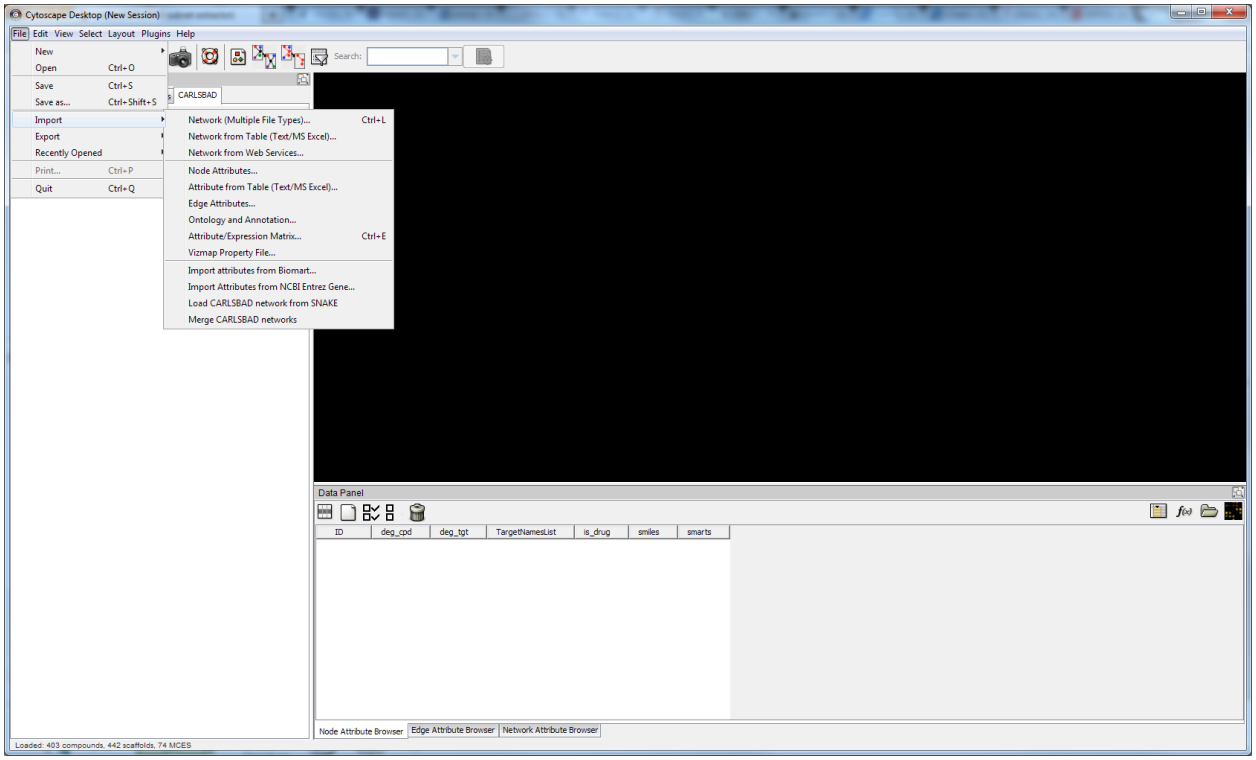
3. **<u>The CARLSBAD Plugin for Cytoscape</u>** - Cytoscape is an open source bioinformatics software platform that CARLSBAD is able to utilize, via the plugin, for powerful network visualization. Combining these capabilities with CARLSBAD's powerful cheminformatics and pattern recognition algorithms will allow researchers to generate hypotheses involving the relationships between biological targets, chemical compounds and their common chemical patterns (CCPs). This section will focus on how to import a file downloaded from SNAKE into the CARLSBAD plugin for Cytoscape, as well as familiarize you with the tools that CARLSBAD provides for network visualization and analysis. It will do this by taking you through a tutorial example that, if you wish to follow, will require you to download the human Caspase-8 target file from SNAKE (see screenshot on previous page).

    - **Importing a file downloaded from SNAKE into the CARLSBAD plugin for Cytoscape -** Once you have performed a query with SNAKE, you can import the file you downloaded into the CARLSBAD plugin for Cytoscape. However before importing, you should familiarize yourself with the location of the CARLSBAD tools available in Cytoscape. This is a simple task as there are two menus you will use for accessing the powerful functions of the CARLSBAD plugin. CARLSBADs analysis tools can be found in the "Plugins" menu at the top of the Cytoscape window as depicted here:




*(Note: More discussion of how to use these tools and what they do can be found in the "Using the CARLSBAD plugin options" section of this manual).*

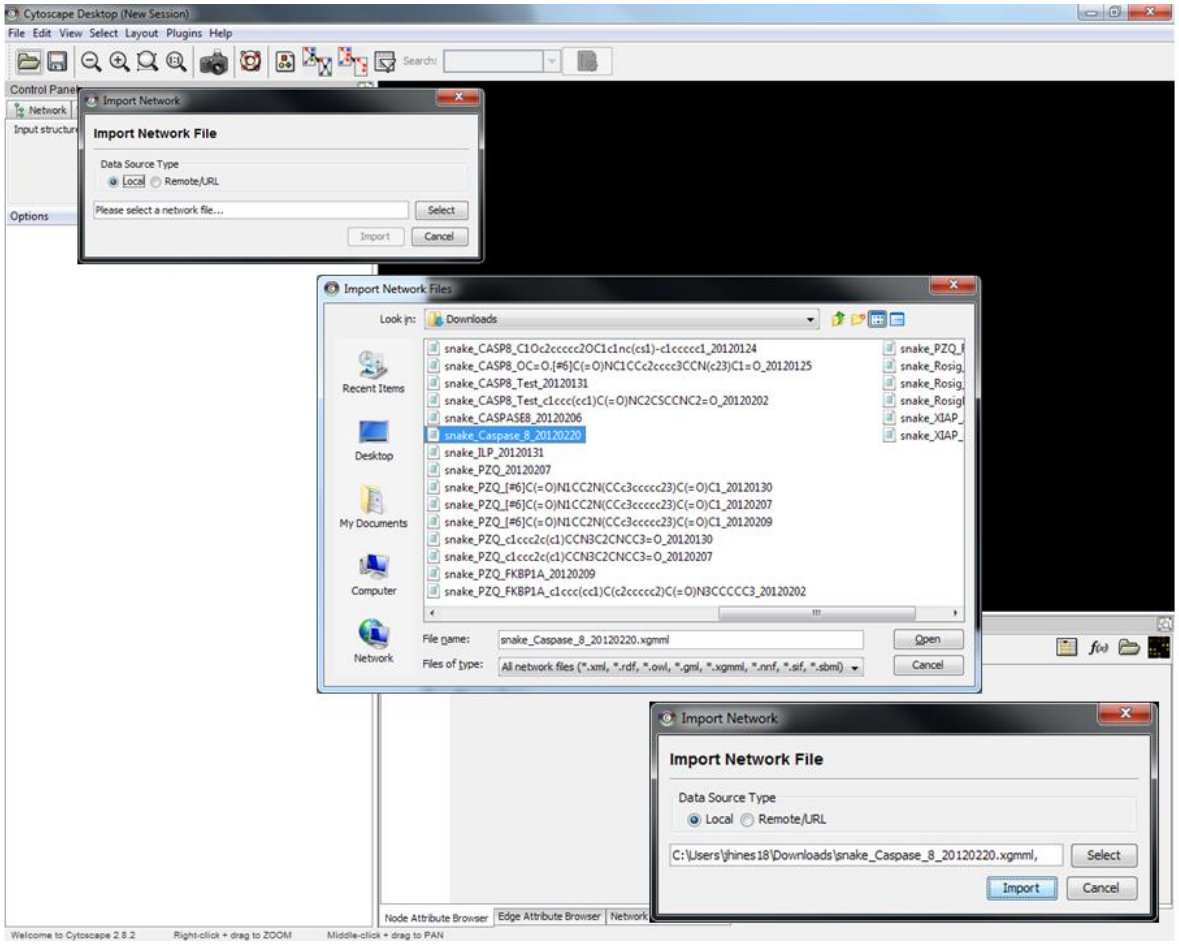
In addition, CARLSBADs import tools can be found in the "File - Import" menu at the top of the Cytoscape window as depicted here:




As you can see, the first of the two import tool options is, “Load CARLSBAD network from SNAKE.” After selecting this option ...

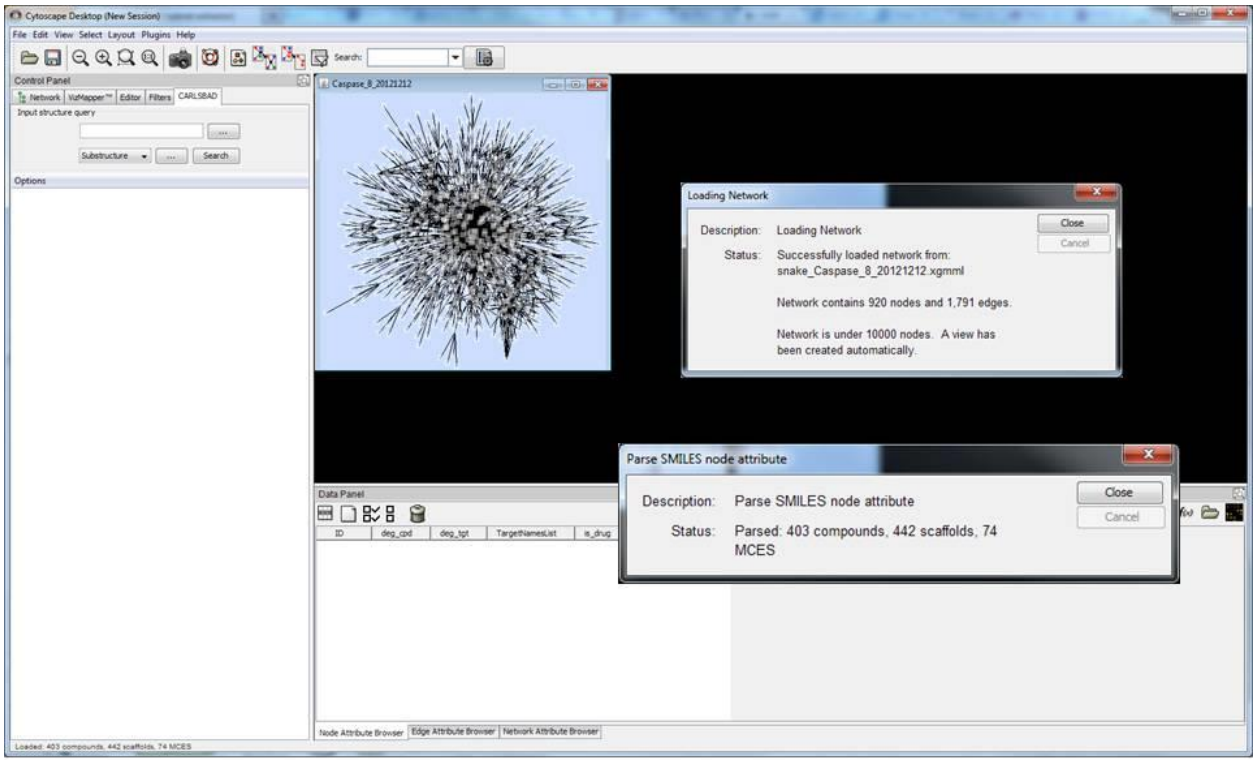
*(Note: More discussion of how to use the second option in this menu, “Merge CARLSBAD networks,” can be found in the "Merging CARLSBAD Networks" section of this manual).*

... the “Import Network” window (top left) will appear. By clicking the “Select” button, the “Import Network Files” window (middle) will appear allowing you to chose the file you downloaded from SNAKE. Finally, after selecting your file, you will see the “Import Network” window (bottom right) again with the file extension. Now you can click “Import” and …



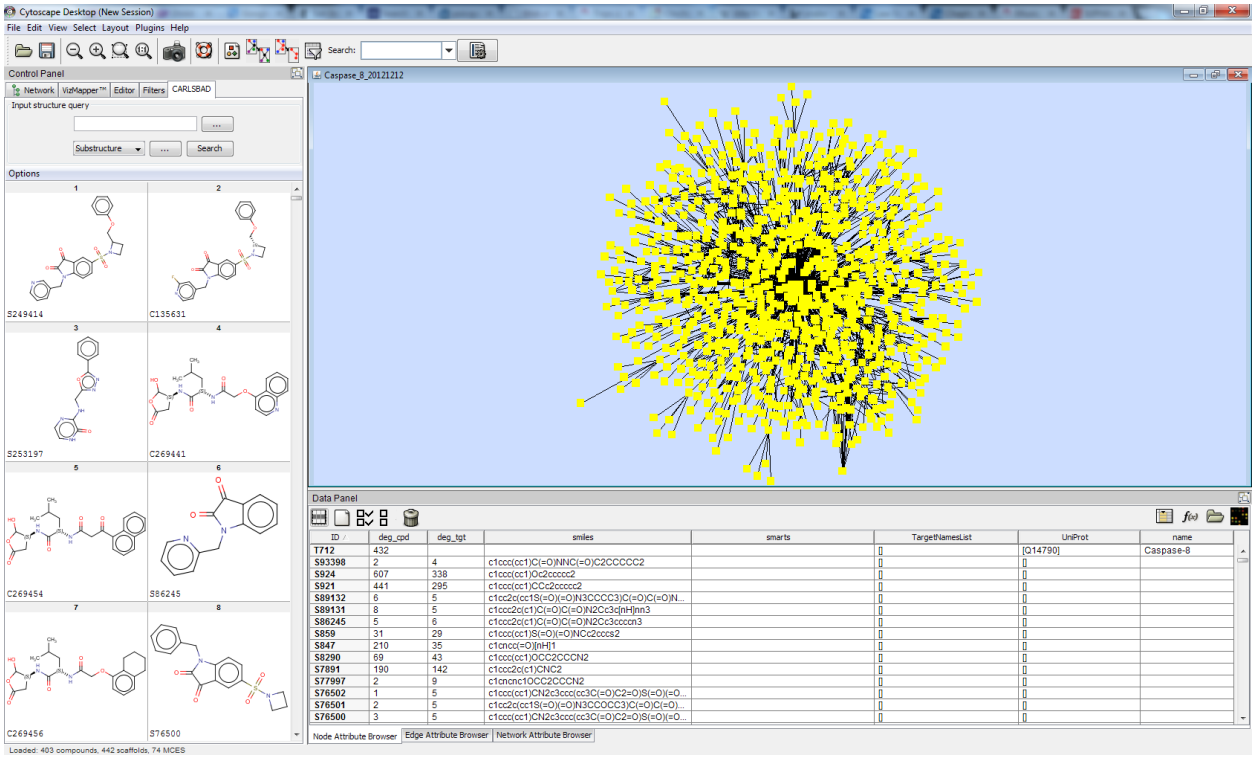

*(Tip: As you can see here, naming your file in SNAKE is important to keeping your network files organized as you might be downloading multiple files from SNAKE. In addition, you must remember to name your file before hitting the “search” button in order for your filename to appear in the file you download).*

… the “Loading Network” window (top) will appear during the importing process. Select “Close” and the import process will finish and display the “Parse SMILES node attribute” window (bottom).


Loading Network
Description: Loading Network
Status: Successfully loaded network from: snake_Caspase_8_20121212.xgmml
Network contains 920 nodes and 1,791 edges.
Network is under 10000 nodes. A view has been created automatically.
Close
Cancel
Parse SMILES node attribute
Description: Parse SMILES node attribute
Status: Parsed: 403 compounds, 442 scaffolds, 74 MCES
Close
Cancel

*(Tip: After closing the “Parse SMILES node attribute” window, you will likely want to maximize the network view by clicking the square “expand window” button above the network depicted and then selecting the “1:1 Magnifying Glass” button highlighted at the top of the screen in order to maximize and center the network view).*

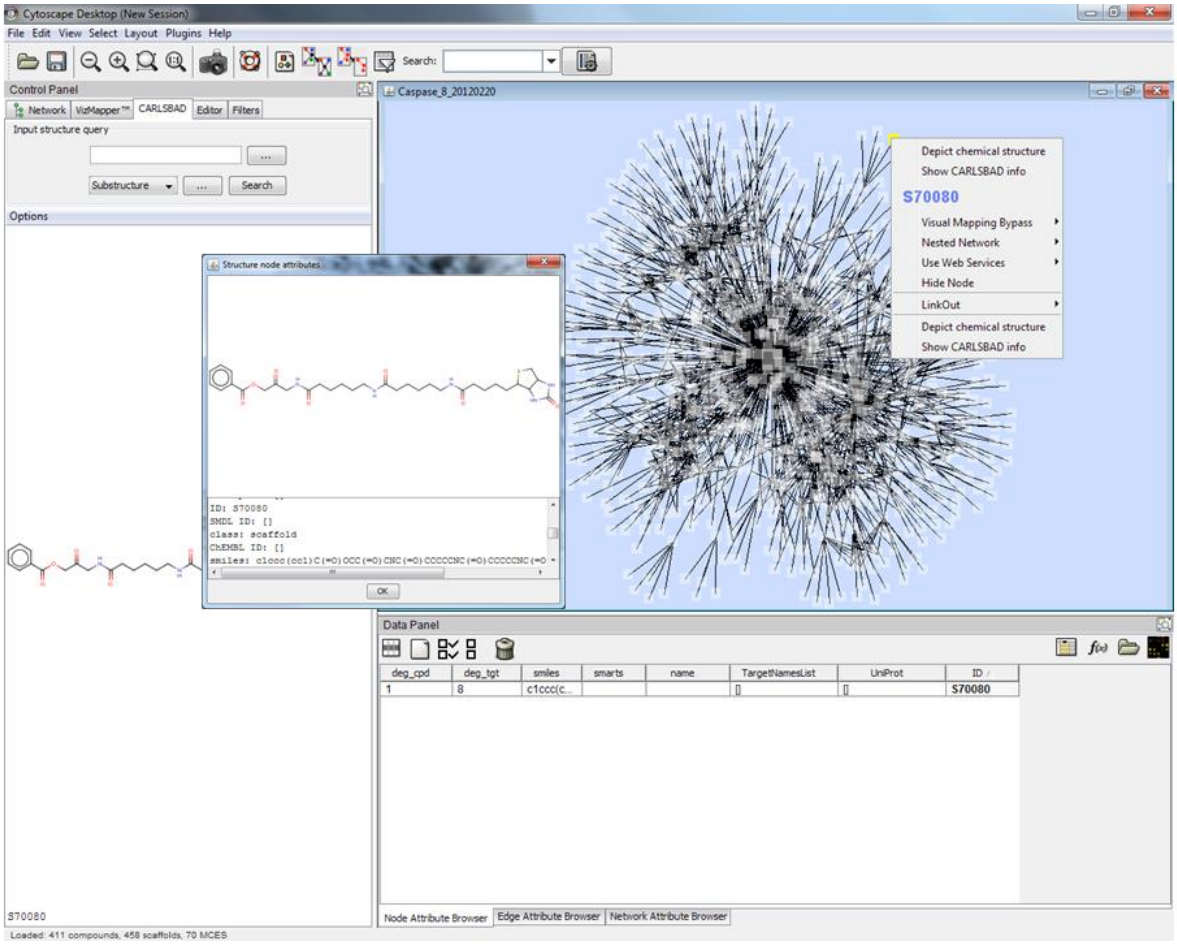
- **CARLSBAD Node Attributes and Cytoscape Viewing Panels -** There are three panels in Cytoscape: the "**Control Panel**" (left), the "**Data Panel**" (bottom) and the "**Network View Panel**" (main window depicted with network containing yellow nodes below). In the "Network View Panel," you can select all nodes by either hitting “Ctrl A” or going to the “Select” drop-down menu at the top and choosing “Nodes.” This latter menu option provides for further node selection capabilities.




After selecting “some” or “all” nodes, you will see structural depiction in the “Control Panel” (left) and node attributes in the “Data Panel” (bottom).

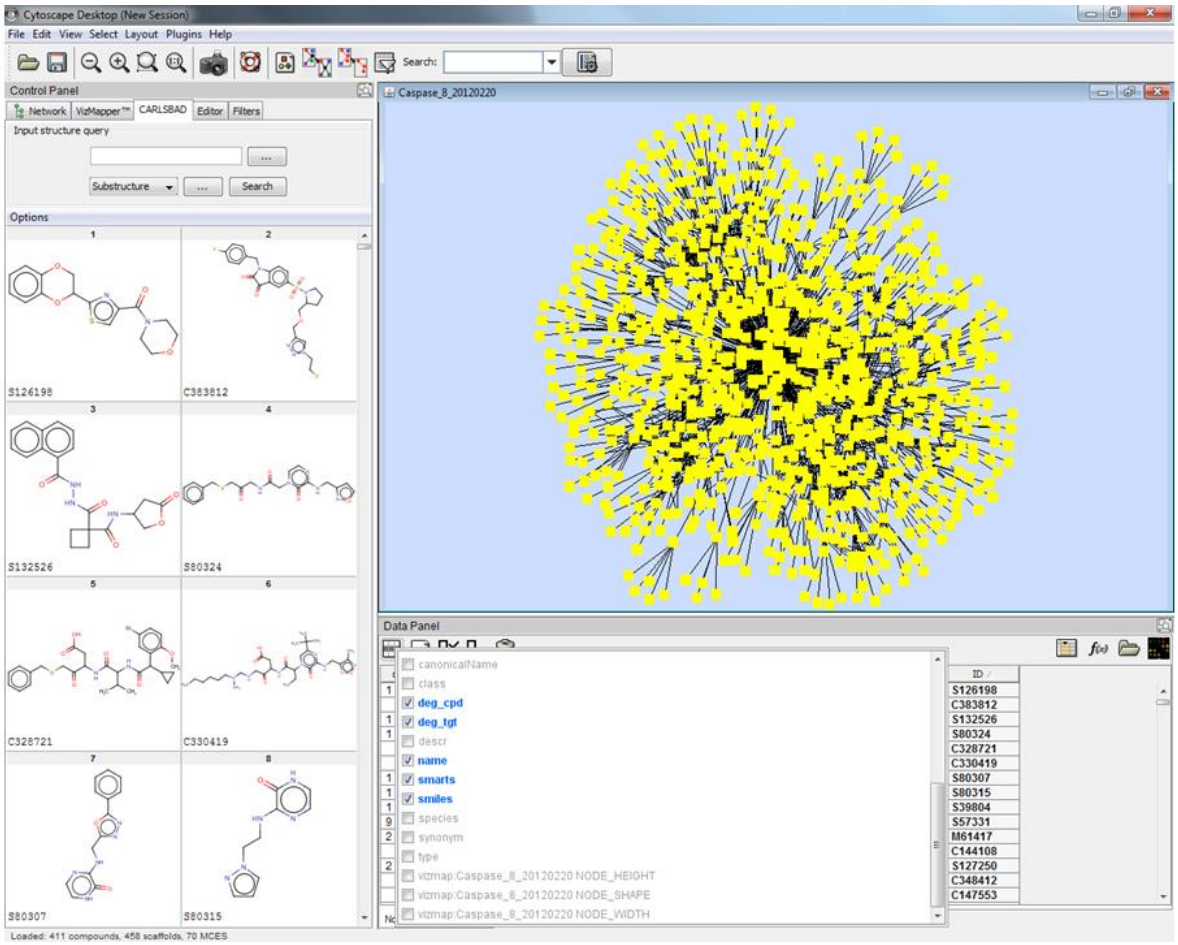
*(Note: In Cytoscape target nodes are represented by octagons, compounds by squares, scaffolds by diamonds and MCESs by hexagons).*

The attributes in the "Data Panel" that load automatically after importing a file are a subset of all the CARLSBAD attributes contained in the file you downloaded from SNAKE. The entire set of attributes can be accessed in two ways. To access all the attribute information of a specific node, you can right click on a node in the “Network View Panel” and select the “Show CARLSBAD info” option, which will then reveal a pop-up window containing all the attribute information for that node. In addition ...

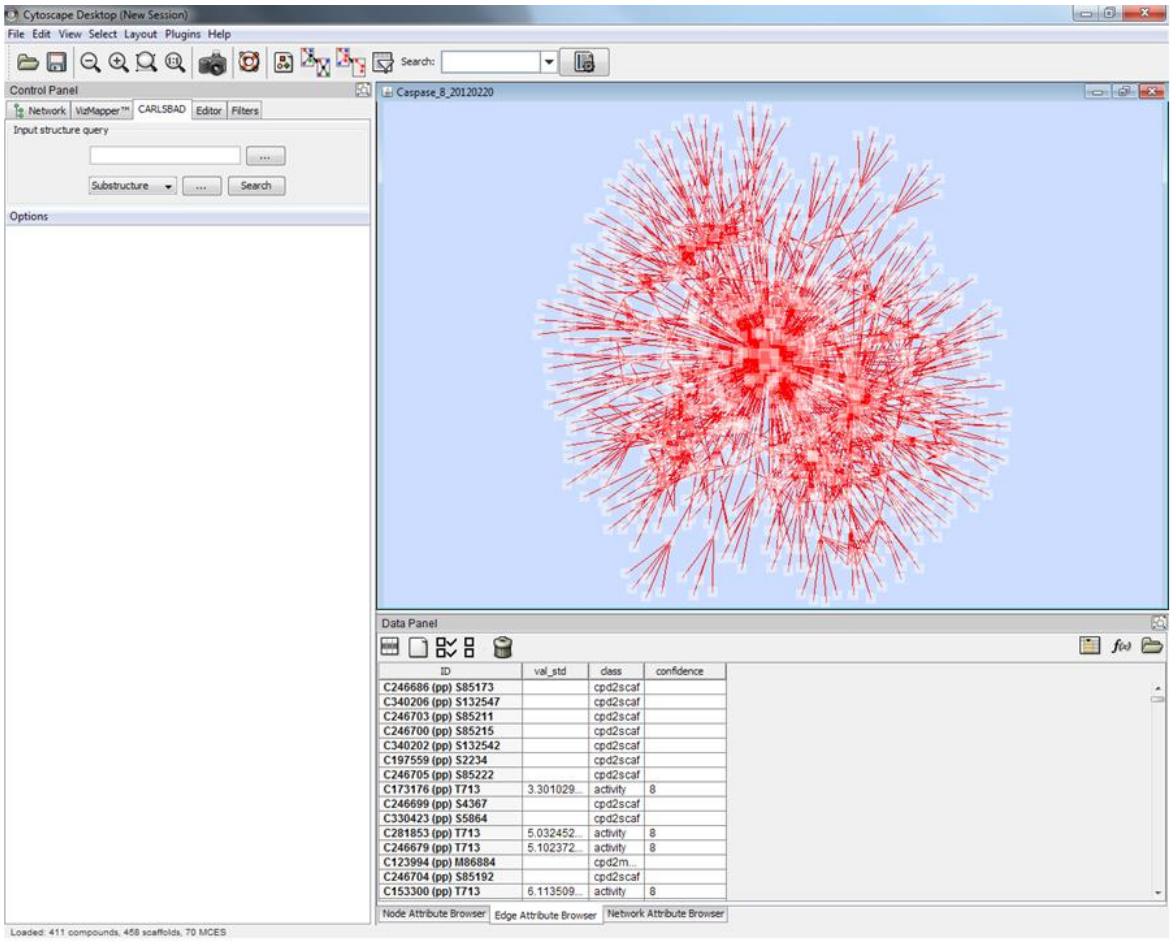



*(Note: This section is designed to help you get comfortable navigating the CARLSBAD plugin to Cytoscape. In this section you will find where and what attributes are included in the CARLSBAD platform; i.e., bioactivities, common identifiers, SMILES, structures, etc. However, a more detailed explanation of individual attributes can be found in the "Glossary" at the end of this manual).*

.... you can obtain all the attribute information for all nodes (or a subset of nodes) in the "Data Panel" by selecting the “Grid” button in the top left of the “Data Panel” and then checking the attributes you wish to display (depicted below). Also, you can simply check the “2X Box and 2X Check Mark” button to the right of the “Grid” button, which will display all the attribute information in the “Data Panel.”



- **CARLSBAD Edge Attributes** - So far we have been learning about accessing node attributes, however, important information such as bioactivity is contained in the **edges**. To access edge attributes, you will need to select edges and then chose the "**Edge Attribute Browser**" tab at the bottom of the "Data Panel" next to the "**Node Attribute Browser**" tab that has been selected in the previous screenshots in this section of the manual. You can select all edges by either hitting “Alt A” or going to the “Select” drop-down menu at the top and choosing “Edges.” This latter menu option provides for further edge selection capabilities.

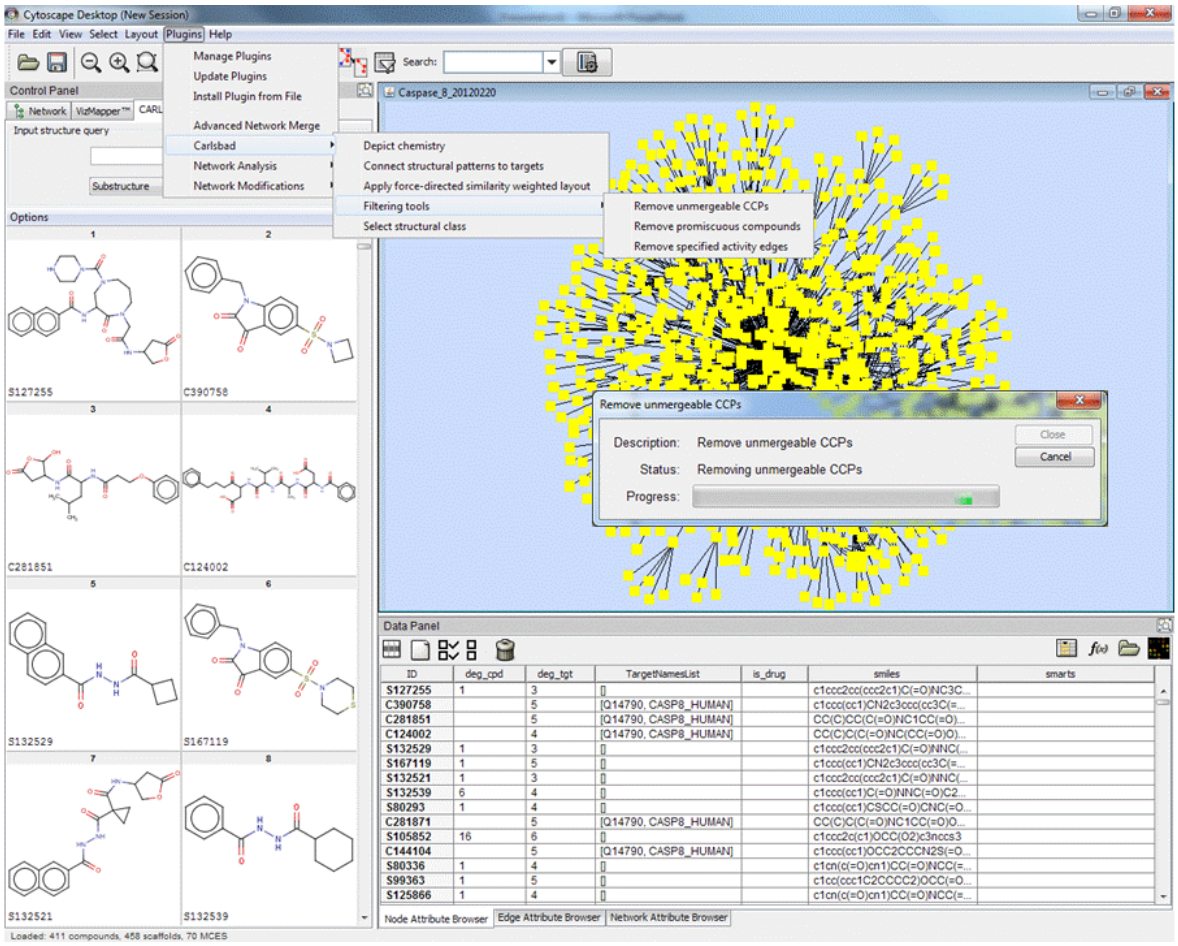



After selecting “some” or “all” edges, you will see edge attributes in the “Data Panel” at the bottom (please see the “Glossary” at the end of this manual for specific definitions).

*(Note: Like the node attributes, the SNAKE file you downloaded contains additional edge attributes that can be seen by the same methods as were previously described for viewing additional node attributes).*

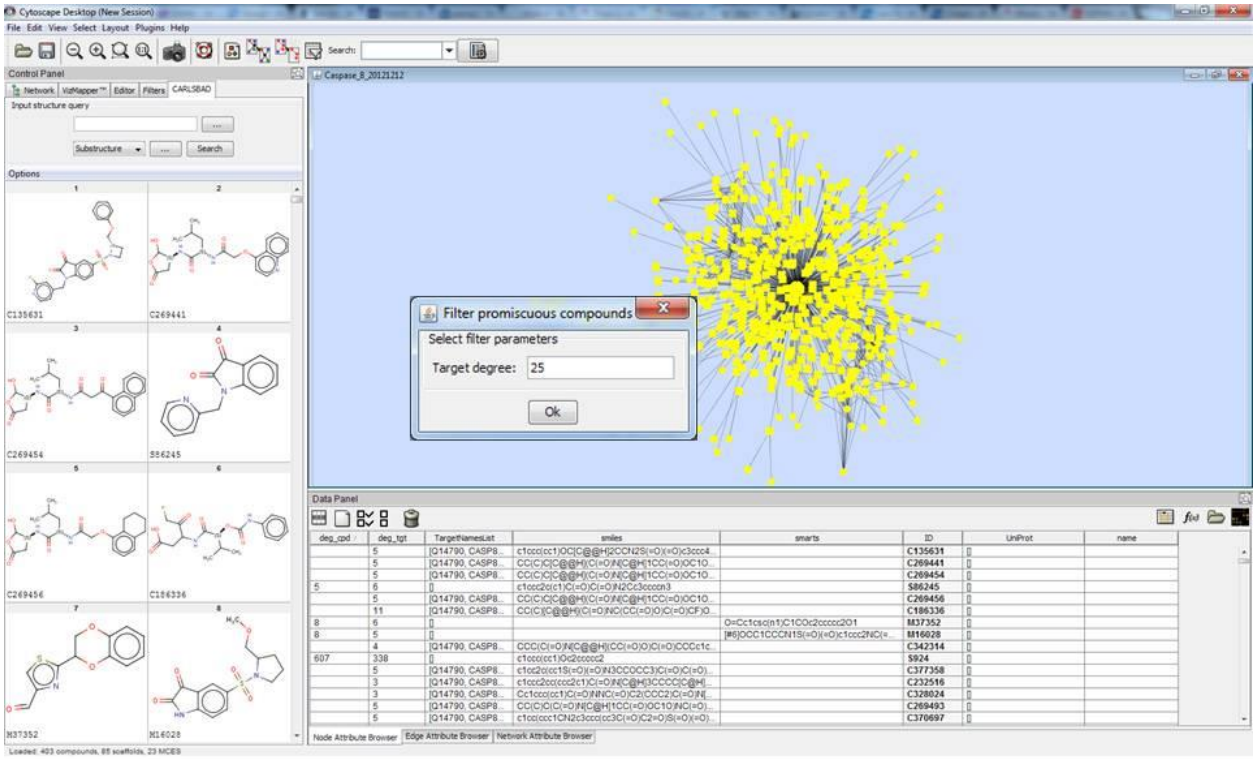
- **Using the CARLSBAD plugin options: Filtering tools -** In the CARLSBAD plugin you will find three "Filtering Tools" that can help you refine your network analysis. These filters can be used separately or in combination. *However, it is important to note that if using filters in combination, you should apply the filters in the order they appear.*

    *i.* **Remove unmergeable CCPs** - This filter removes **scaffolds** and **MCESs** that are only associated with compounds in the **current network**. That means the CCPs removed have no association with any other compounds from the CARLSBAD database and therefore cannot be used to find new potential **lead compounds** or **repurposed drugs** for your target(s) of interest.



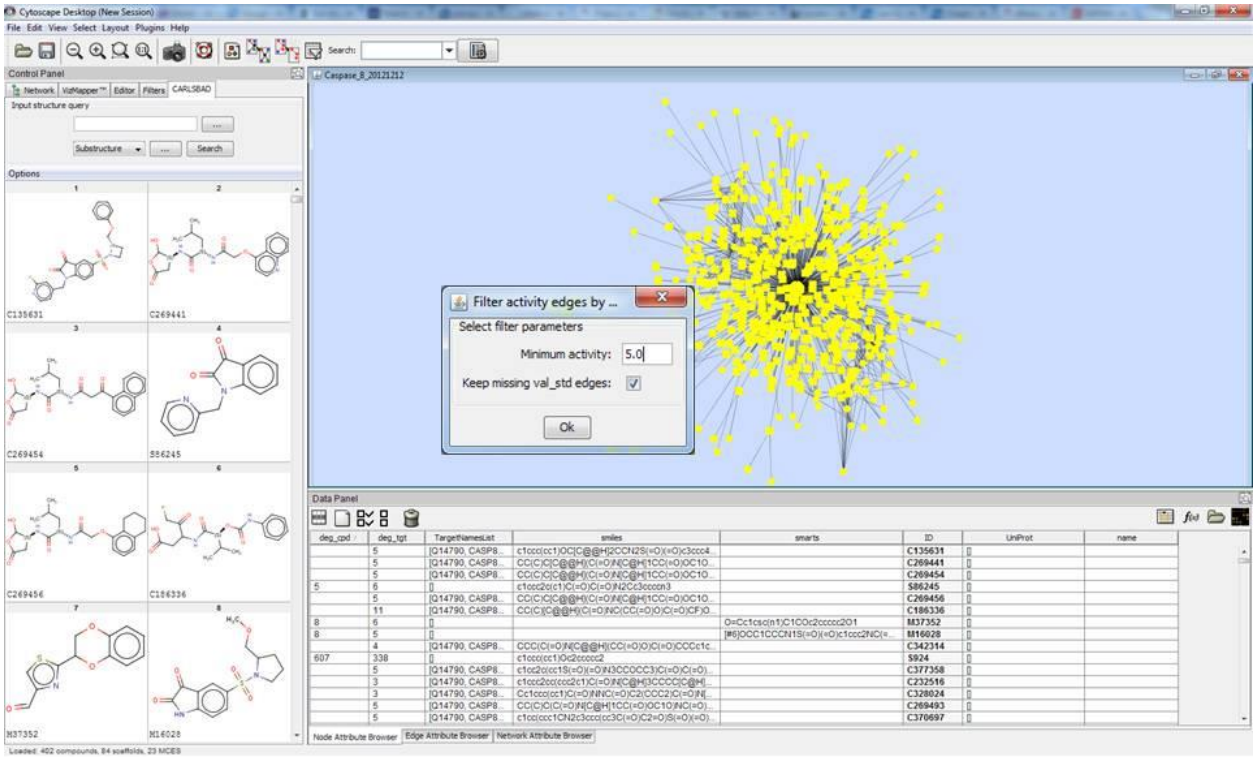

*(Note: To improve the speed of this filter, ensure you don't have any nodes selected when applying the filter as Cytoscape will attempt to continuously regenerate structural depictions in the "Control Panel." If this happens it will not harm anything, but it may take a few moments longer to process).*

ii. **Remove promiscuous compounds** - This filter provides a user defined option for the "Target Degree," an attribute providing information as to the number of targets a CCP or compound relates to in the entire CARLSBAD database (as opposed to just the targets in the file downloaded from SNAKE).




*(Note: You may notice that this filter can remove compounds and CCPs despite the fact the filter says it's for "Promiscuous compounds." This is because CCPs that only associate with compound(s) that are removed will also be removed. This is a convenient feature of this filter (as well as the next filter we will use, "Remove specified activity edges") because it not only removes those nodes which were defined by the filter, but it also removes the "**disconnected nodes**," or those CCPs that only connected to the target through compounds which were removed).*

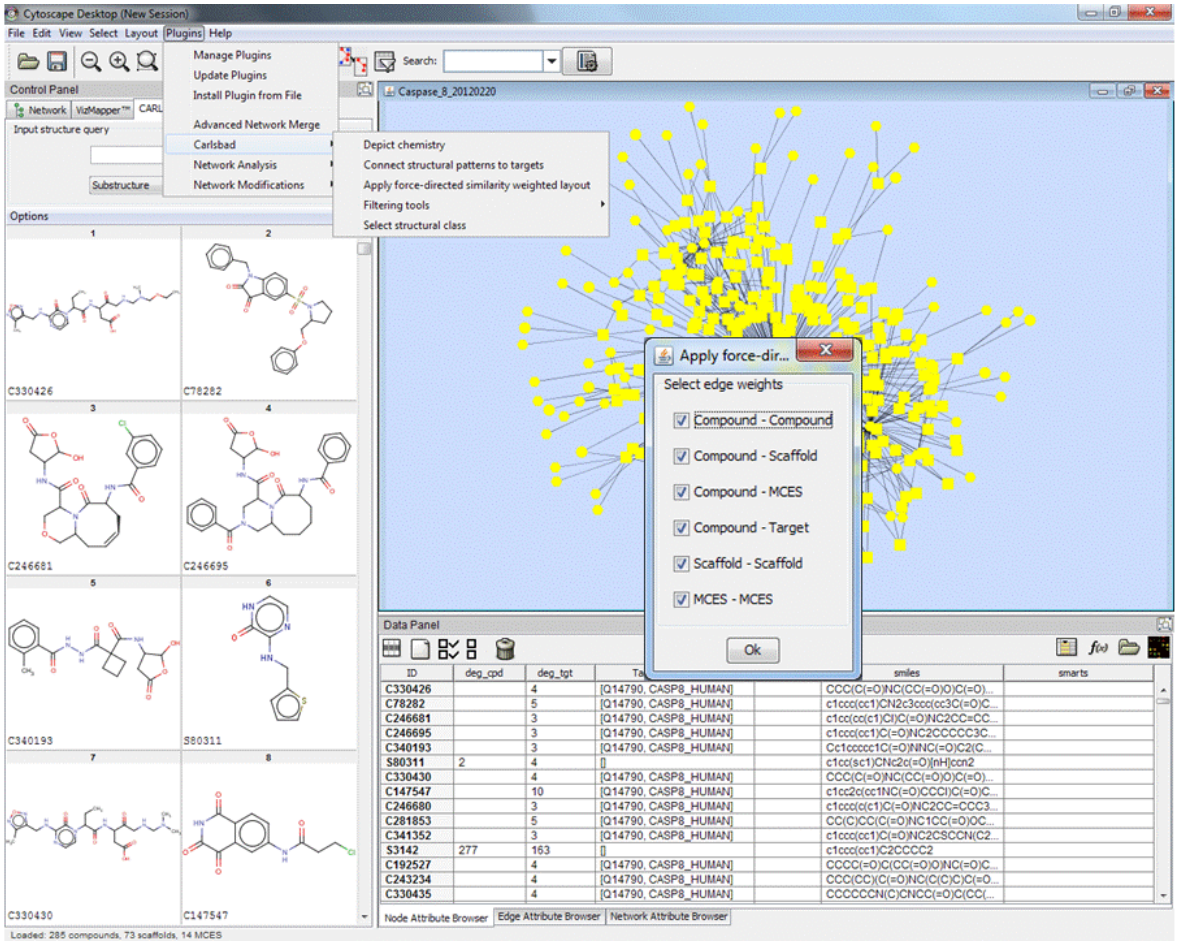
iii. **Remove specified activity edges** - This filter provides user defined options such as minimum activity, which is a bioactivity attribute (converted to –log molar so that below 5.0 is the same as bioactivities above 10 uM). In addition, there is a box that is checked. On occasion a compound may not have a specific known bioactivity for the target and by checking this box (the default selection) those compounds will remain in your network.




*(Note: To improve the speed of this filter, ensure you don't have any nodes selected when applying the filter as Cytoscape will attempt to continuously regenerate structural depictions in the "Control Panel." If this happens it will not harm anything, but it may take a few moments longer to process).*

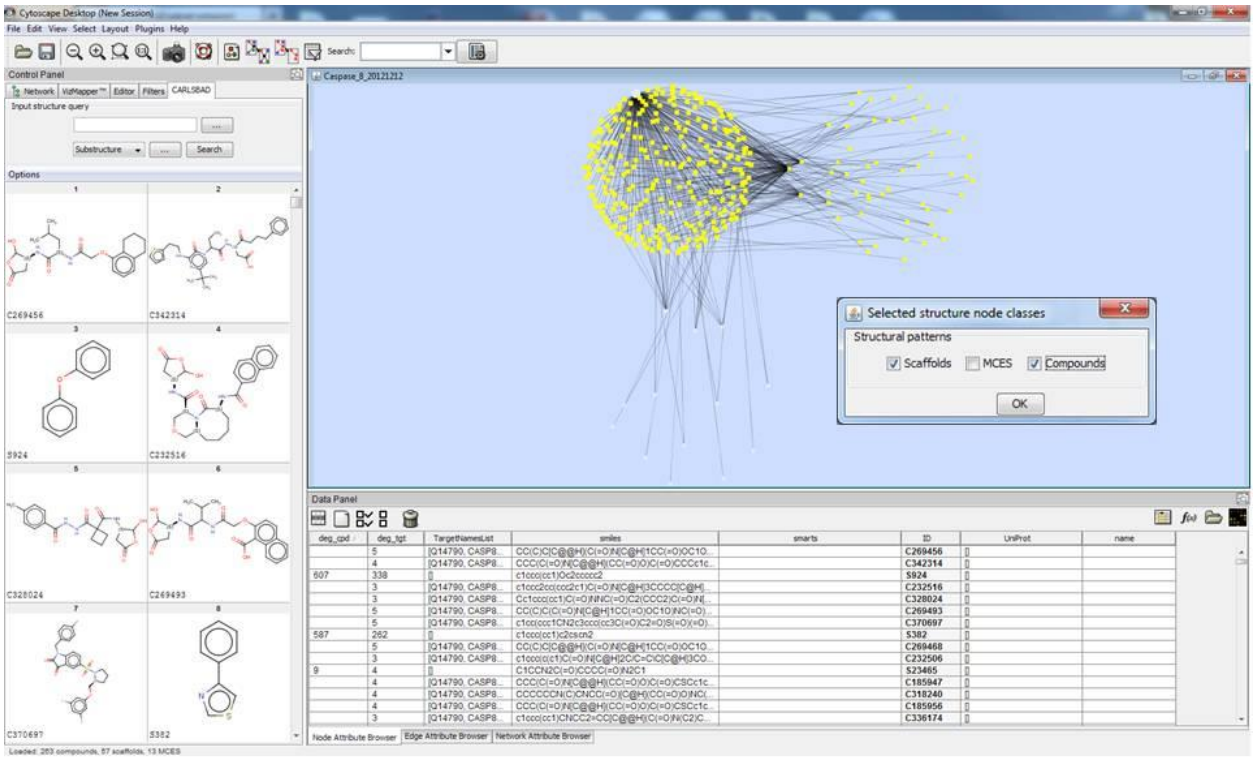
- **Using the CARLSBAD plugin options: Additional tools -** In the CARLSBAD plugin you will find some additional tools that CARLSBAD offers to assist with your analysis in addition to the "Filtering Tools" discussed in the previous section.

    i. **Apply force-directed similarity weighted layout** - The CARLSBAD plugin provides the ability to weigh the network visualization based on node and edge similarities such as chemical patterns and activity values. After selecting the “Apply force-directed similarity weighted layout” option, a window will appear allowing you to select which attribute relationship(s) you would like to favor when rearranging your network. All choices refer to structures, so that nodes will cluster based on how structurally similar they are to each other, except for the "Compound - Target" option, which refers to bioactivities such that the greater the affinity a compound has for a target, the closer it will be to that target after the network rearrangement.



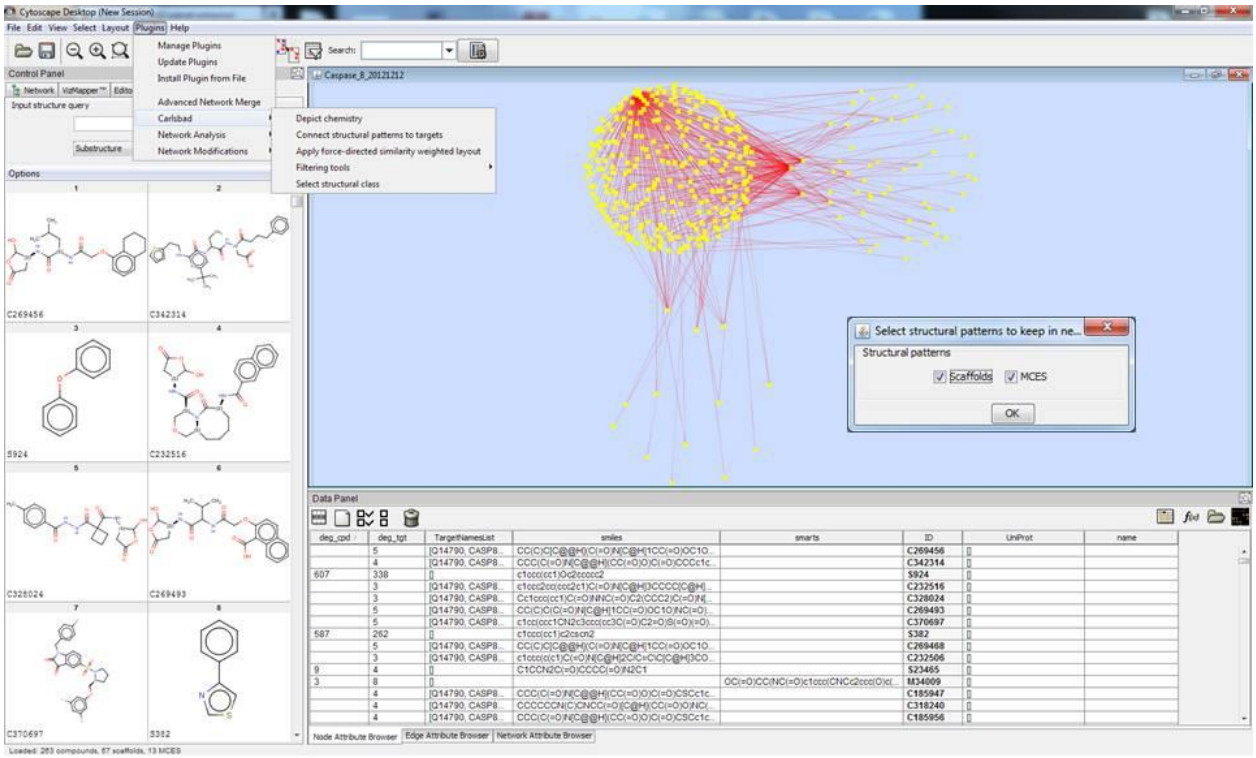

*(Note: This filter may take several minutes to compute and while it is doing so you will see new edges drawn and subsequently erased).*

ii. **Select structural class** - Now that your network has been rearranged with “Apply force-directed similarity weighted layout," you might want to use the “Select structural class” tool to highlight specific structures. This tool can be used at anytime to select either scaffolds, MCESs or compounds separately or in combination. In the screenshot below, we have chosen to select scaffolds and compounds to highlight how the previous tool, “Apply force-directed similarity weighted layout,” rearranged the network.



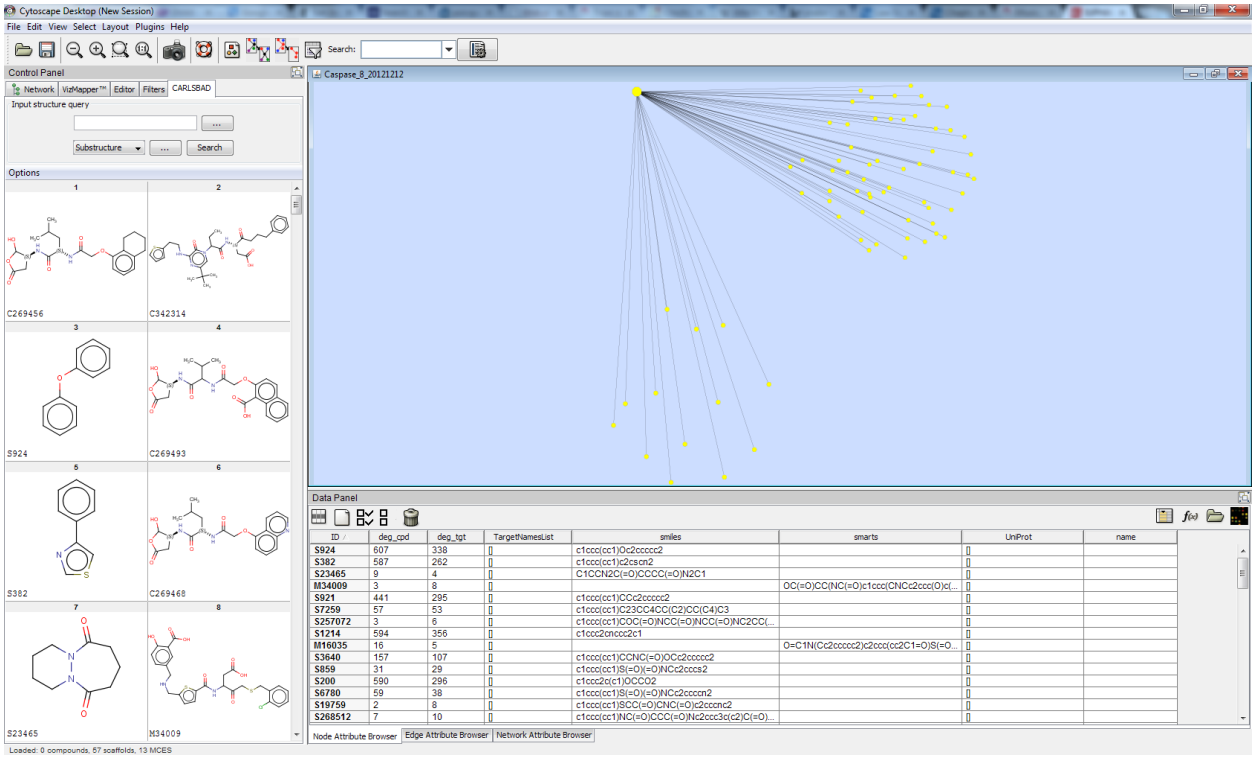

As you can see, the nodes clustered in yellow on the right are scaffolds, the nodes clustered in yellow in the center are the compounds and the nodes that remain unhighlighted at the bottom are MCESs.

iii. **Connect structural patterns to targets** - This tool will remove the compounds in your network and connect the CCPs directly to the target. When using “Connect structural patterns to targets” you will lose some attribute information such as the activity values (since those values depict a relationship between the target and the compounds, the latter of which you are removing with this filter). However, if you have already applied the filter for activity (see the "Filtering tools" section), you will have some idea as to the lowest bioactivity that could be left in your network. In addition, with the “Connect structural patterns to targets” tool, you have the ability to choose either Scaffolds, MCESs or both to connect directly to the target(s).




*(Note: To improve the speed of this filter, ensure you don’t have any nodes selected when applying the filter as Cytoscape will attempt to continuously regenerate structural depictions in the “Control Panel.” If this happens it will not harm anything, but it may take a few moments longer to process).*

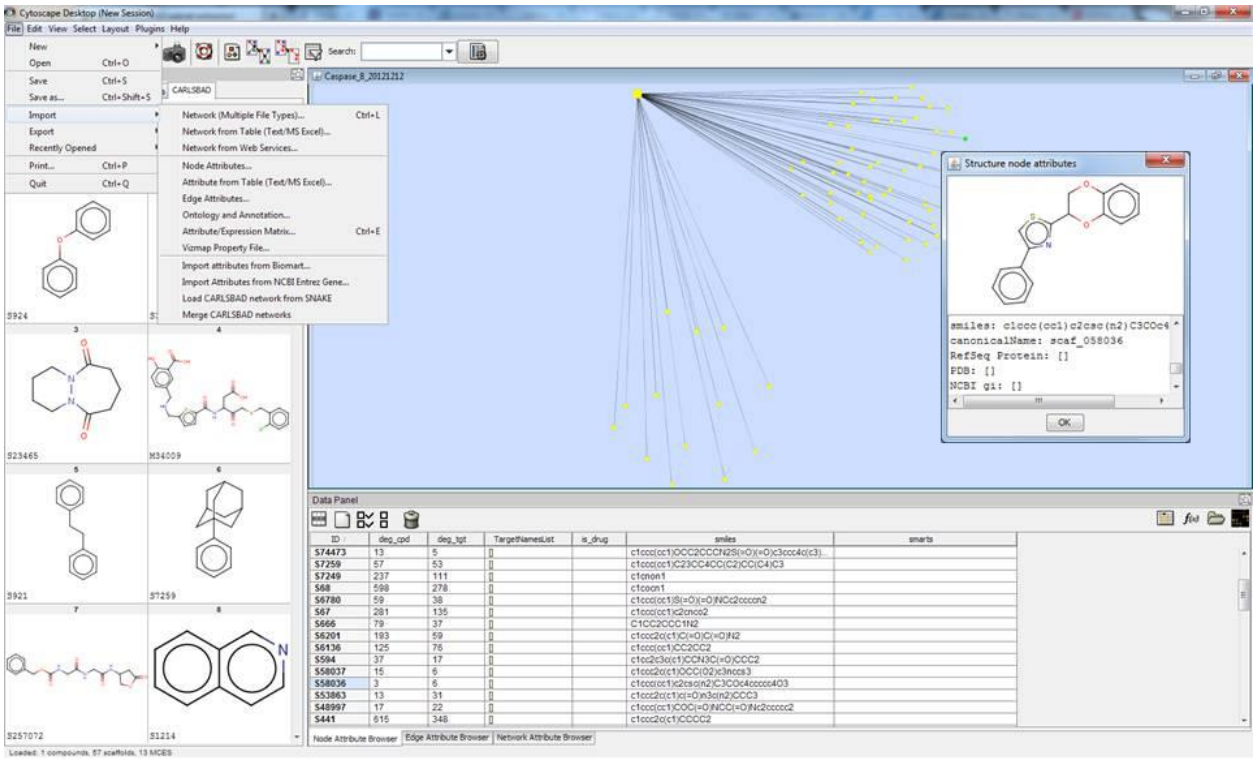
After following the CARLSBAD "Plugin" steps described in this document and applying all the filters and tools described thus far, we have now reduced the original Caspase-8 network, which contained 403 compounds, 442 scaffolds and 74 MCES, down to just 57 scaffolds and 13 MCES that are all associated with compounds which have a minimum activity for the target of 5.0 –log molar, are mergeable (a concept we will explore further in the next section) and are not associated with compounds that are “too promiscuous” by our own definition...




...In addition, we have removed the compounds in the network and arranged the nodes based on attribute similarities. A depiction of this can be seen in the screenshot below and can be compared to the network we began with in the "CARLSBAD Node Attributes and Cytoscape Viewing Panels" section of this manual.

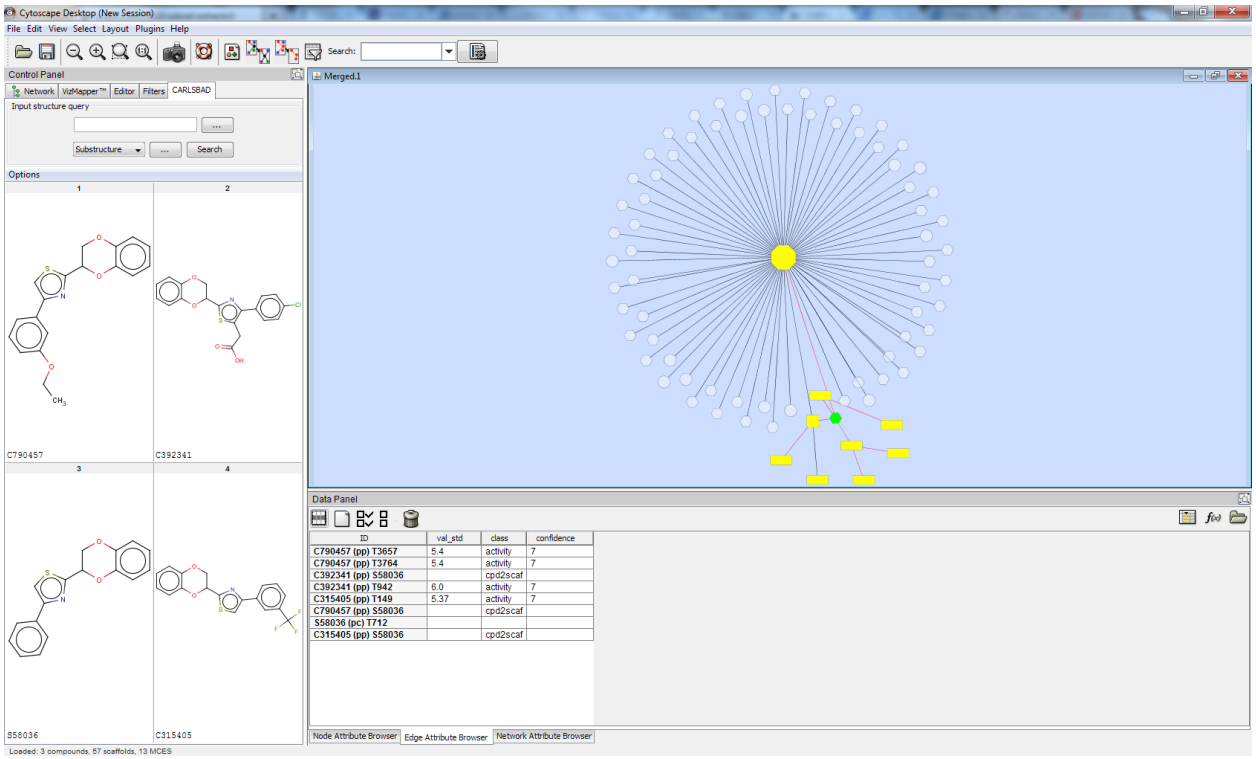
4. **Merging CARLSBAD Networks** - **Network growth**, the ability to grow your current network based on the network of another node that is in your current network, can be achieved in CARLSBAD by the merging of networks downloaded from SNAKE. To do this, we must generate a network file from SNAKE, as we learned in the section, "SNAKE: SubNet Application Kit and Extractor for CARLSBAD," but for a node in the current network for which we are interested. In this manual, our current network is the Caspase-8 network that we have been working with and based on our analysis thus far, we will choose a CCP to merge and thus grow this network.

    - We begin this process by first choosing a CCP (in this case we will be using S58036, which also can be represented with SMILES c1ccc(cc1)c2csc(n2)C3COc4ccccc4O3), then querying SNAKE for the network based file associated with the chosen CCP (by either querying the scaffold ID or using SMILES in the structure box of the compound query in SNAKE, this process is discussed in the "SNAKE: SubNet Application Kit and Extractor for CARLSBAD" section of this manual) and finally downloading that file into Cytoscape with the CARLSBAD plugin by ...




... selecting the "Merge CARLSBAD networks" option from the "File – Import" drop-down menu. Next, we simply select the file as we previously did in the "Importing a file downloaded from SNAKE into the CARLSBAD plugin for Cytoscape" section of this manual.

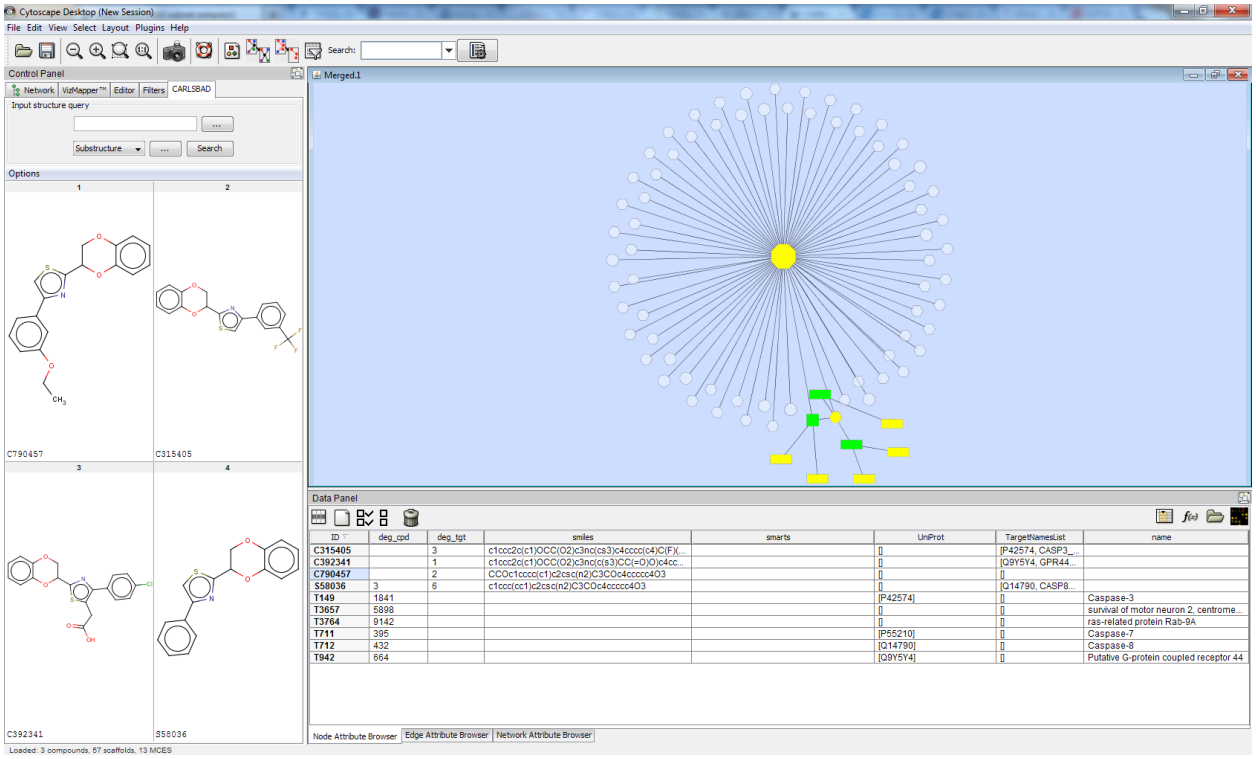
After merging networks, you can see the new growth network highlighted in yellow, the original TOI, Caspase-8, highlighted in yellow (octagon in the middle) and the scaffold we grew highlighted in green.




In addition, by viewing the Data Panel towards the bottom of the screen with the "Edge Attribute Browser" tab selected, we can see that there are four new activities and they are above 5.0 –log molar (remember we lost the activity value between the original scaffolds and Caspase-8 in a previous step when we connected the CCPs to the target by removing the compounds known to associate with Caspase-8 from the CARLSBAD database, but we know that activity is also above 5.0 -log molar since we filtered out all compounds with activities below 5.0). Also …..

*(Note: When growing a network, if a node is a CCP or a compound, be sure to use the appropriate query area in SNAKE. In addition, be aware that you can use node attribute information from your current network to facilitate your search in SNAKE; i.e., SMILES or CARLSBAD ID, however, if using the latter be aware the CARLSBAD IDs change, so it is recommended in your research to track structures by something more static like SMILES).*

….. We can see in the "Data Panel" towards the bottom of the screen with the "Node Attribute Browser" tab selected that there are three new compounds and five new targets that have been grown. In the "Data Panel" you can see the names of all the targets. Also, the new compounds are highlighted in green.




The newly merged compounds are potential lead compounds, however, further research should be conducted before concluding they are definitely leads for the **target of interest**. In addition, it should be recognized that we only grew one of 70 CCPs here, so continuing the merging process as described in this section will likely lead to more potential lead compounds and even some drugs that might be potentially repurposed.

*(Note: Compounds that are known drugs are tagged with the node attribute "is_drug" and can be identified easily in CARLSBAD).*

**This manual and tutorial should provide a researcher with the necessary tools to quickly develop a logical and creative hypothesis. However, further research should be undertaken before beginning experimentation. Some of the steps that might be undertaken are discussed further in the subsequent section, "Next Steps."**

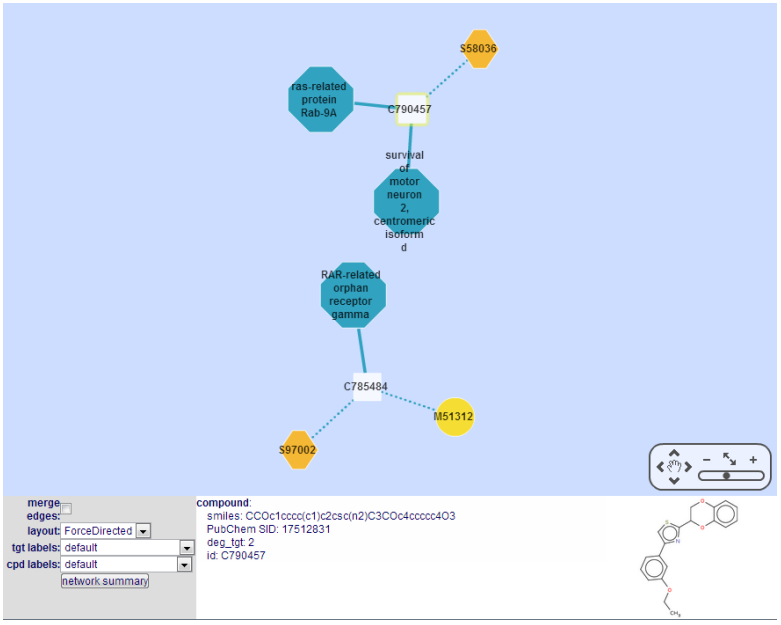
5. **Next Steps** - The CARLSBAD platform provides researchers with the ability to generate novel hypotheses pertaining to the relationships between biological targets, chemical compounds and their common chemical patterns (CCPs). Hypothesis development is a highly personal and creative process beginning with a well defined question. The powerful network theory and algorithms behind CARLSBAD can help guide researchers with a well-defined question towards a hypothesis, but CARLSBAD can only go so far. This section focuses on some steps researchers can take to help transition from their CARLSBAD findings to experimentation.

- **Literature Searches -** After conducting a process of discovering potential new chemical and biological relationships, the researcher is left with the task of confirming that those relationships haven't been previously studied and then clearly defining their own individual hypothesis for testing. CARLSBAD cannot define your hypothesis, however, some steps you might want to consider include:

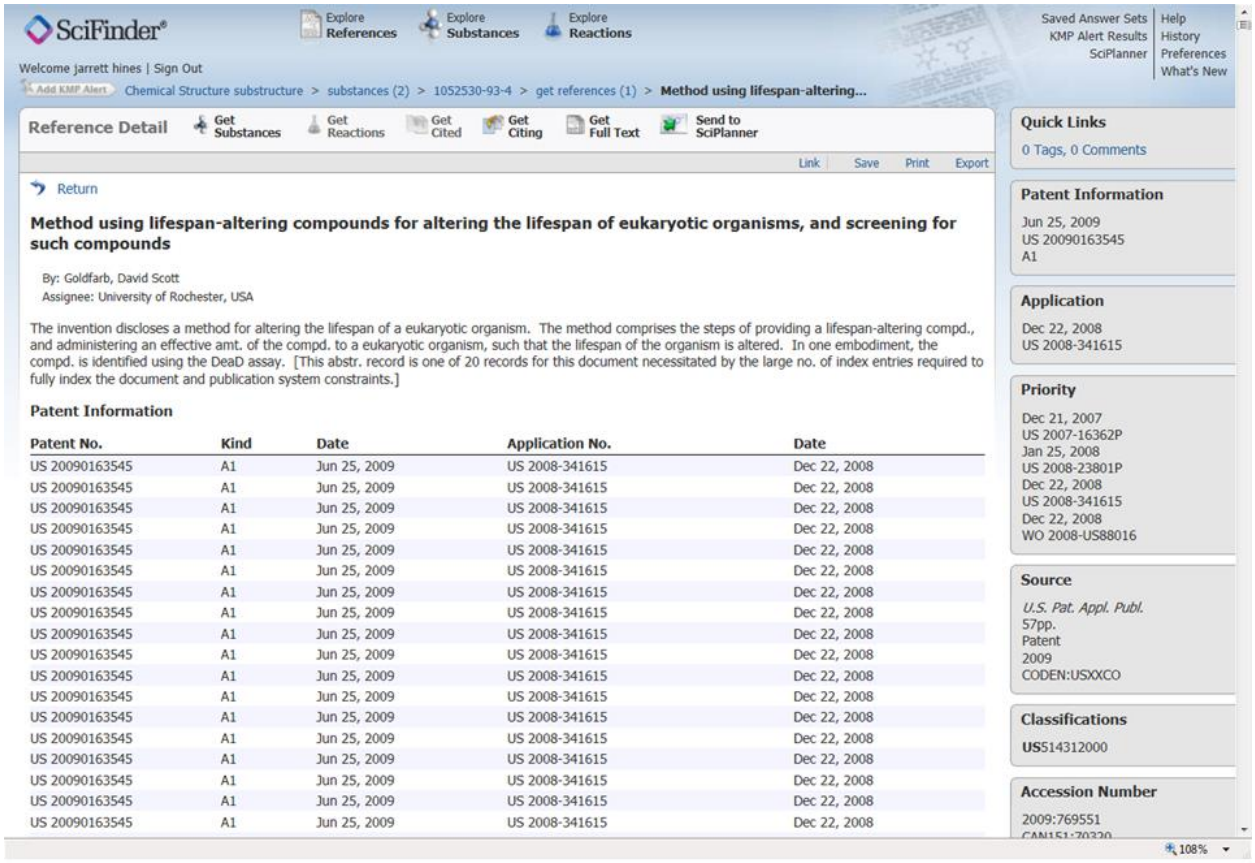
    i. Searching a database like “SciFinder” to see if a possible lead that you identified with CARLSBAD has any literature relating your target and the potential lead compound.

    ii. Searching a database like “Web of Science” to see what relationships may exist between the class of molecule and target for which you are interested.

As was highlighted in the previous section, "Merging CARLSBAD Networks," the merging of networks generated three new compounds for one of the 70 CCPs that we refined the network down to and explored. We are going to concentrate here on the compound with the CARLSBAD ID C790457 and SMILES CCOc1cccc(c1)c2csc(n2)C3COc4ccccc4O3.

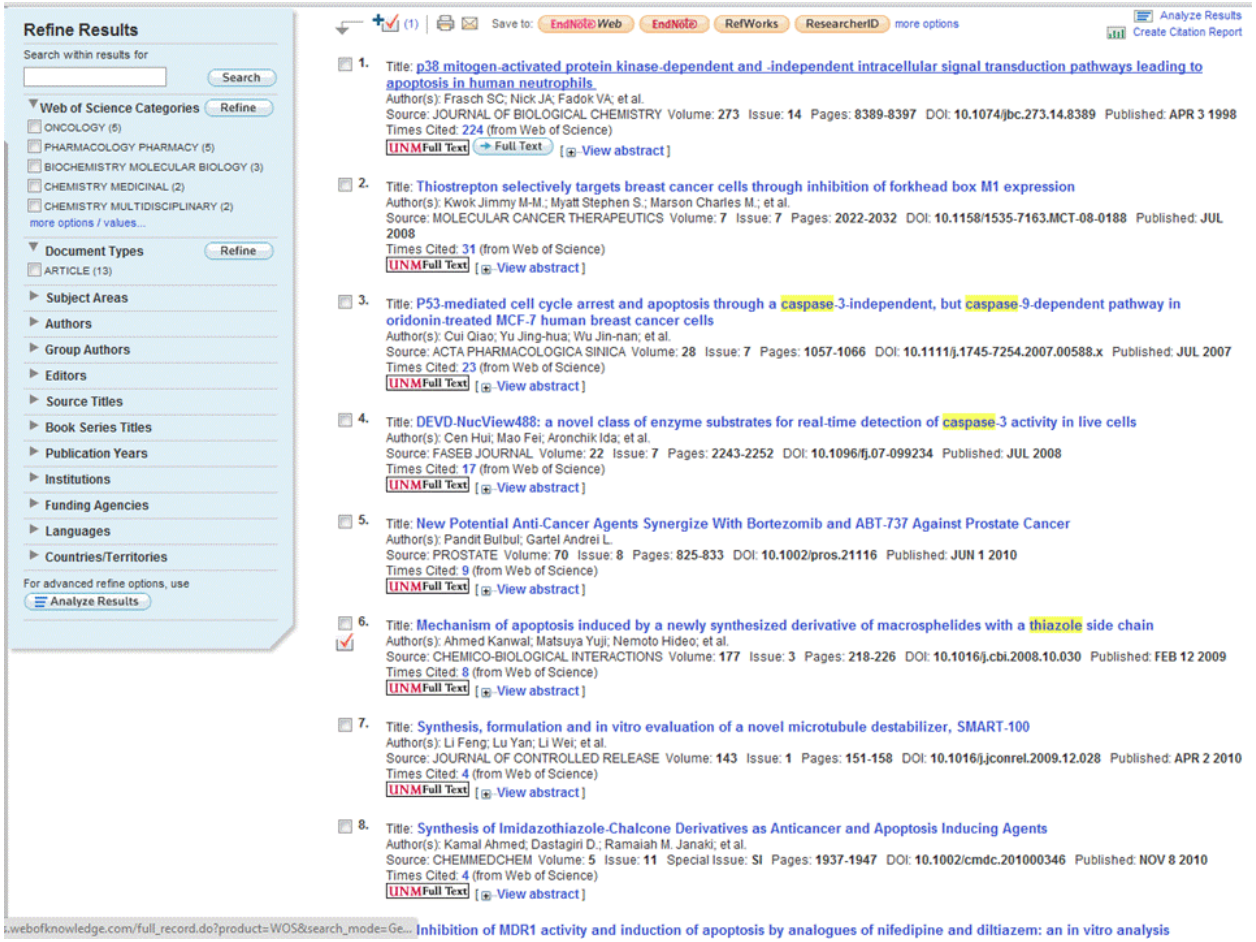


CARLSBAD ID C790457 was identified as 2-(2,3-dihydro-1,4-benzodioxin-2-yl)-4-(3-ethoxyphenyl)-thiazole in SciFInder. We can see there is only one reference to the molecule and it is for a patent that was filed in 2009 pertaining to “lifespan-altering” compounds. This is interesting since we know our target, Caspase-8, plays both a critical role in apoptosis and autophagy, two processes that could be directly linked to the aging process.


SciFinder
Explore References
Explore Substances
Explore Reactions
Saved Answer Sets
KMP Alert Results
SciPlanner
Help
History
Preferences
What's New
Welcome Jarrett hines | Sign Out
Chemical Structure substructure > substances (2) > 1052530-93-4 > get references (1) > Method using lifespan-altering...
Reference Detail
Get Substances
Get Reactions
Get Cited
Get Citing
Get Full Text
Send to SciPlanner
Link
Save
Print
Export
Return
Method using lifespan-altering compounds for altering the lifespan of eukaryotic organisms, and screening for such compounds
By: Goldfarb, David Scott
Assignee: University of Rochester, USA
The invention discloses a method for altering the lifespan of a eukaryotic organism. The method comprises the steps of providing a lifespan-altering compd., and administering an effective amt. of the compd. to a eukaryotic organism, such that the lifespan of the organism is altered. In one embodiment, the compd. is identified using the DeaD assay. [This abstr. record is one of 20 records for this document necessitated by the large no. of index entries required to fully index the document and publication system constraints.]
Patent Information
Patent No. Kind Date Application No. Date
US 20090163545 A1 Jun 25, 2009 US 2008-341615 Dec 22, 2008
Quick Links
0 Tags, 0 Comments
Patent Information
Jun 25, 2009
US 20090163545
A1
Application
Dec 22, 2008
US 2008-341615
Priority
Dec 21, 2007
US 2007-16362P
Jan 25, 2008
US 2008-23801P
Dec 22, 2008
US 2008-341615
Dec 22, 2008
WO 2008-US88016
Source
U.S. Pat. Appl. Publ.
57pp.
Patent
2009
CODEN:USXXCO
Classifications
US514312000
Accession Number
2009:769551
108%

*(Note: CARLSBAD IDs change, so it is recommended in your research to track structures by something more static like SMILES/SMARTS).*

Finally, performing a literature search for, "thiazole and caspase," in “Web of Science” reveals 13 articles, one of which looks specifically at Caspase-8 (depicted as the 6th result in the screenshot below) and none of which involves the potential lead compound identified in this manual.

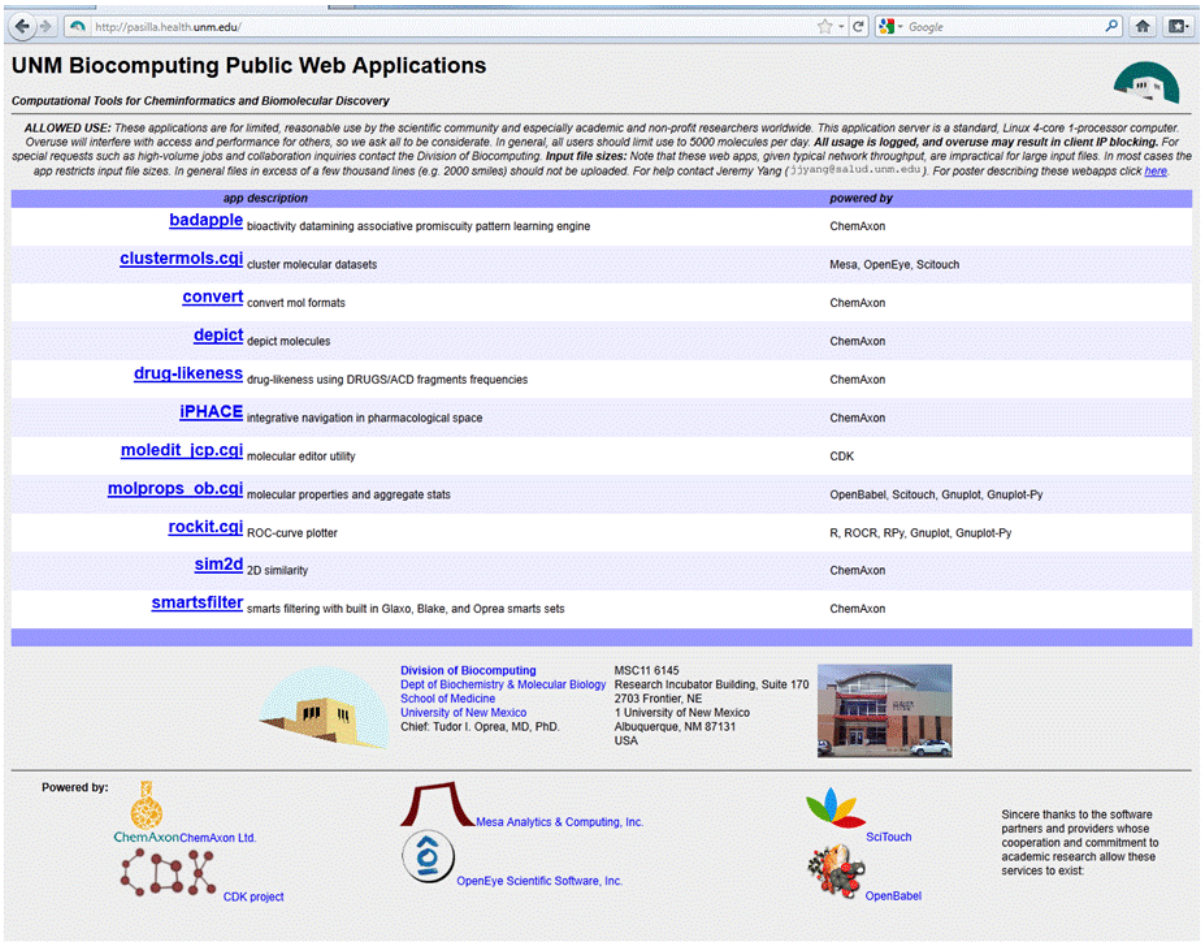

Refine Results
Search within results for
Search
Web of Science Categories
Refine
ONCOLOGY (5)
PHARMACOLOGY PHARMACY (5)
BIOCHEMISTRY MOLECULAR BIOLOGY (3)
CHEMISTRY MEDICINAL (2)
CHEMISTRY MULTIDISCIPLINARY (2)
more options / values...
Document Types
Refine
ARTICLE (13)
Subject Areas
Authors
Group Authors
Editors
Source Titles
Book Series Titles
Publication Years
Institutions
Funding Agencies
Languages
Countries/Territories
For advanced refine options, use
Analyze Results
Save to: EndNote Web EndNote RefWorks ResearcherID more options
Analyze Results
Create Citation Report
1. Title: p38 mitogen-activated protein kinase-dependent and -independent intracellular signal transduction pathways leading to apoptosis in human neutrophils
Author(s): Frasch SC; Nick JA; Fadok VA; et al.
Source: JOURNAL OF BIOLOGICAL CHEMISTRY Volume: 273 Issue: 14 Pages: 8389-8397 DOI: 10.1074/jbc.273.14.8389 Published: APR 3 1998
Times Cited: 224 (from Web of Science)
UNM Full Text Full Text [ View abstract ]
2. Title: Thiostrepton selectively targets breast cancer cells through inhibition of forkhead box M1 expression
Author(s): Kwok Jimmy M-M.; Myatt Stephen S.; Marson Charles M.; et al.
Source: MOLECULAR CANCER THERAPEUTICS Volume: 7 Issue: 7 Pages: 2022-2032 DOI: 10.1158/1535-7163.MCT-08-0188 Published: JUL 2008
Times Cited: 31 (from Web of Science)
UNM Full Text [ View abstract ]
3. Title: P53-mediated cell cycle arrest and apoptosis through a caspase-3-independent, but caspase-9-dependent pathway in oridonin-treated MCF-7 human breast cancer cells
Author(s): Cui Qiao; Yu Jing-hua; Wu Jin-nan; et al.
Source: ACTA PHARMACOLOGICA SINICA Volume: 28 Issue: 7 Pages: 1057-1066 DOI: 10.1111/j.1745-7254.2007.00588.x Published: JUL 2007
Times Cited: 23 (from Web of Science)
UNM Full Text [ View abstract ]
4. Title: DEVD-NucView488: a novel class of enzyme substrates for real-time detection of caspase-3 activity in live cells
Author(s): Cen Hui; Mao Fei; Aronchik Ida; et al.
Source: FASEB JOURNAL Volume: 22 Issue: 7 Pages: 2243-2252 DOI: 10.1096/fj.07-099234 Published: JUL 2008
Times Cited: 17 (from Web of Science)
UNM Full Text [ View abstract ]
5. Title: New Potential Anti-Cancer Agents Synergize With Bortezomib and ABT-737 Against Prostate Cancer
Author(s): Pandit Bulbul; Gartel Andrei L.
Source: PROSTATE Volume: 70 Issue: 8 Pages: 825-833 DOI: 10.1002/pros.21116 Published: JUN 1 2010
Times Cited: 9 (from Web of Science)
UNM Full Text [ View abstract ]
6. Title: Mechanism of apoptosis induced by a newly synthesized derivative of macrosphelides with a thiazole side chain
Author(s): Ahmed Kanwal; Matsuya Yuji; Nemoto Hideo; et al.
Source: CHEMICO-BIOLOGICAL INTERACTIONS Volume: 177 Issue: 3 Pages: 218-226 DOI: 10.1016/j.cbi.2008.10.030 Published: FEB 12 2009
Times Cited: 8 (from Web of Science)
UNM Full Text [ View abstract ]
7. Title: Synthesis, formulation and in vitro evaluation of a novel microtubule destabilizer, SMART-100
Author(s): Li Feng; Lu Yan; Li Wei; et al.
Source: JOURNAL OF CONTROLLED RELEASE Volume: 143 Issue: 1 Pages: 151-158 DOI: 10.1016/j.jconrel.2009.12.028 Published: APR 2 2010
Times Cited: 4 (from Web of Science)
UNM Full Text [ View abstract ]
8. Title: Synthesis of Imidazothiazole-Chalcone Derivatives as Anticancer and Apoptosis Inducing Agents
Author(s): Kamal Ahmed; Dastagiri D.; Ramaiah M. Janaki; et al.
Source: CHEMMEDCHEM Volume: 5 Issue: 11 Special Issue: SI Pages: 1937-1947 DOI: 10.1002/cmdc.201000346 Published: NOV 8 2010
Times Cited: 4 (from Web of Science)
UNM Full Text [ View abstract ]
i.webofknowledge.com/full_record.do?product=WOS&search_mode=Ge...
Inhibition of MDR1 activity and induction of apoptosis by analogues of nifedipine and diltiazem: an in vitro analysis

Thus, we might hypothesize that, 2-(2,3-dihydro-1,4-benzodioxin-2-yl)-4-(3-ethoxyphenyl)-thiazole, is a potential lead compound for Caspase-8.

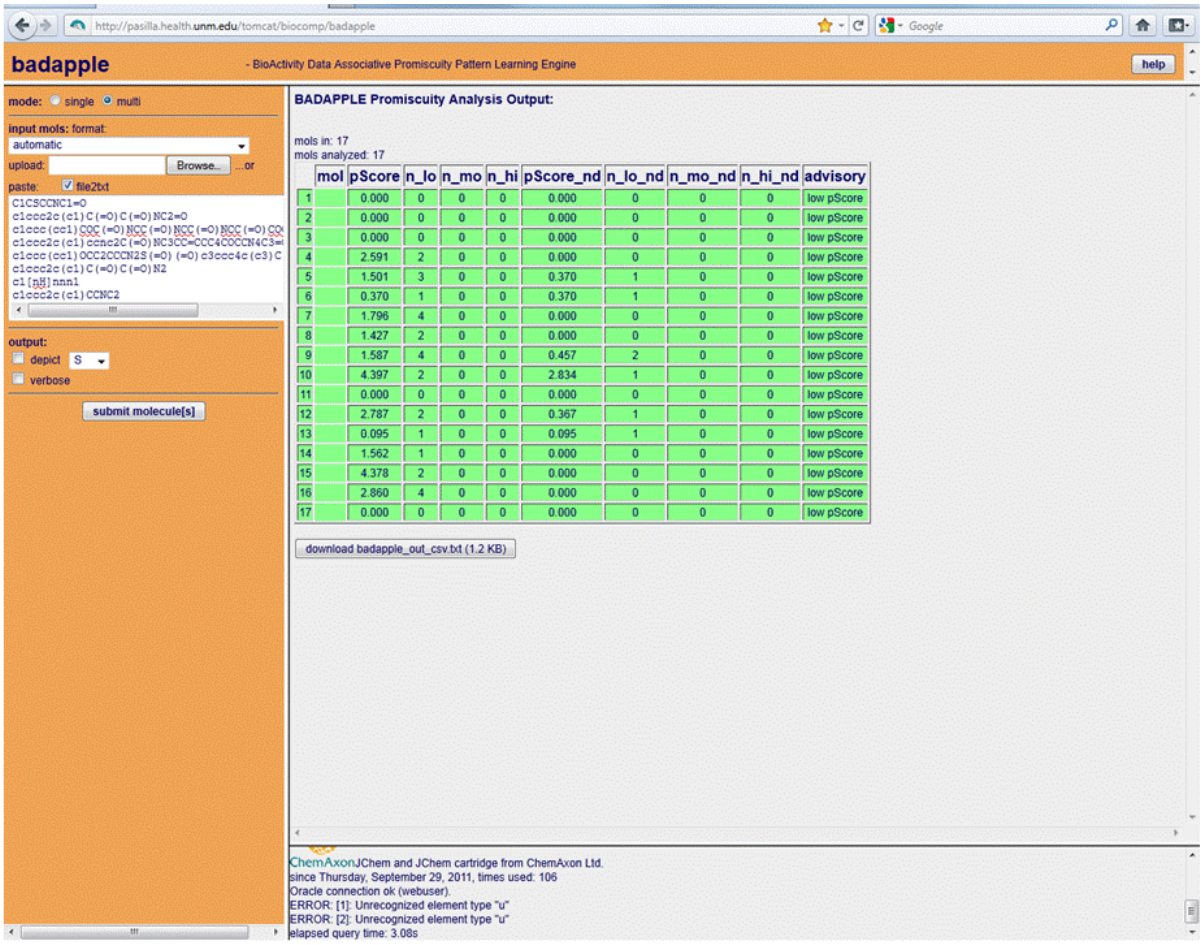
*(Tip: You can use the attribute, “is_drug,” to determine if a compound is a known drug, thus turning your lead compound hypothesis into a potential for drug repurposing).*

- **Additional Biocomputing Tools for Filtering** - In the previous part of this section, "Literature Searches," we analyzed the growth of one CCP, but in reality a researcher would likely go through the process of merging more CCPs into the original network to get a larger list of potential lead compounds. However, this might yield hundreds or even thousands of potential leads.

- To assist with logically reducing your lead compound list further, there are two tools developed by UNM's Translational Informatics Division that can help refine your results based on promiscuity and reactivity, called "badapple" and "smartsfilter" respectively. These tools can be found in "UNM Biocomputing Public Web Applications" at **pasilla.helath.unm.edu** or by going to the "CARLSBAD Community" link on the CARLSBAD website **(carlsbad.health.unm.edu)** and looking under "Other Resources."

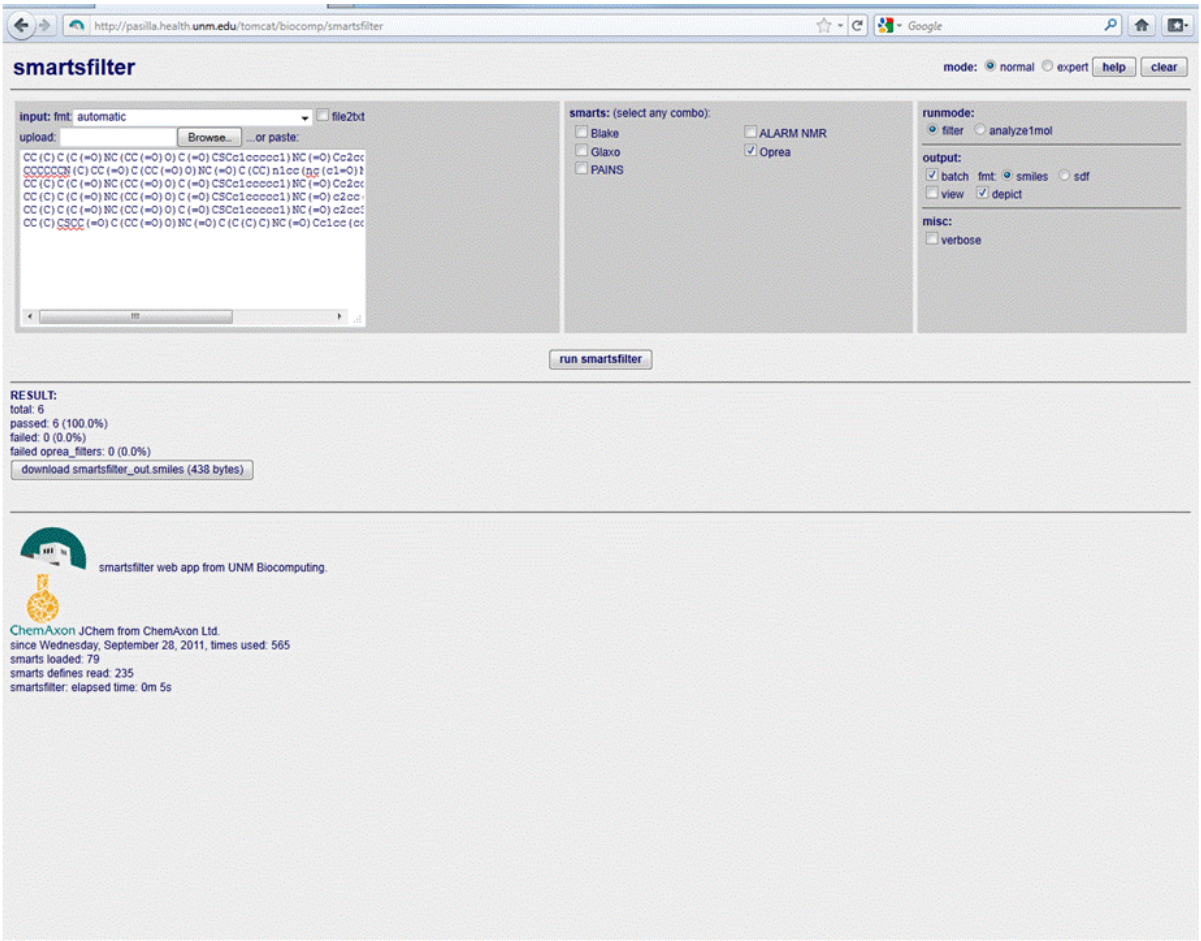

http://pasilla.health.unm.edu/
UNM Biocomputing Public Web Applications
Computational Tools for Cheminformatics and Biomolecular Discovery
ALLOWED USE: These applications are for limited, reasonable use by the scientific community and especially academic and non-profit researchers worldwide. This application server is a standard, Linux 4-core 1-processor computer. Overuse will interfere with access and performance for others, so we ask all to be considerate. In general, all users should limit use to 5000 molecules per day. All usage is logged, and overuse may result in client IP blocking. For special requests such as high-volume jobs and collaboration inquiries contact the Division of Biocomputing. Input file sizes: Note that these web apps, given typical network throughput, are impractical for large input files. In most cases the app restricts input file sizes. In general files in excess of a few thousand lines (e.g. 2000 smiles) should not be uploaded. For help contact Jeremy Yang (jyang@salud.unm.edu). For poster describing these webapps click here.
app description
powered by
badapple bioactivity datamining associative promiscuity pattern learning engine
ChemAxon
clustermols.cgi cluster molecular datasets
Mesa, OpenEye, Scitouch
convert convert mol formats
ChemAxon
depict depict molecules
ChemAxon
drug-likeness drug-likeness using DRUGS/ACD fragments frequencies
ChemAxon
iPHACE integrative navigation in pharmacological space
ChemAxon
moledit_jcp.cgi molecular editor utility
CDK
molprops_ob.cgi molecular properties and aggregate stats
OpenBabel, Scitouch, Gnuplot, Gnuplot-Py
rockit.cgi ROC-curve plotter
R, ROCR, RPy, Gnuplot, Gnuplot-Py
sim2d 2D similarity
ChemAxon
smartsfilter smarts filtering with built in Glaxo, Blake, and Oprea smarts sets
ChemAxon
Division of Biocomputing
Dept of Biochemistry & Molecular Biology
School of Medicine
University of New Mexico
Chief: Tudor I. Oprea, MD, PhD.
MSC11 6145
Research Incubator Building, Suite 170
2703 Frontier, NE
1 University of New Mexico
Albuquerque, NM 87131
USA
Powered by:
ChemAxonChemAxon Ltd.
CDK project
Mesa Analytics & Computing, Inc.
OpenEye Scientific Software, Inc.
SciTouch
OpenBabel
Sincere thanks to the software partners and providers whose cooperation and commitment to academic research allow these services to exist:

i. **BadApple** - This tool will help provide additional promiscuity filtering. Simply paste one or more SMILES directly from the CARLSBAD plugin into the field on the left of the screen and select the button "submit molecule(s)." The structures that pass the promiscuity filter will be highlighted in green with the lower "pScores" (see values in the second column of output) being less promiscuous.




| | mol | pScore | n_lo | n_mo | n_hi | pScore_nd | n_lo_nd | n_mo_nd | n_hi_nd | advisory |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | | 0.000 | 0 | 0 | 0 | 0.000 | 0 | 0 | 0 | low pScore |
| 2 | | 0.000 | 0 | 0 | 0 | 0.000 | 0 | 0 | 0 | low pScore |
| 3 | | 0.000 | 0 | 0 | 0 | 0.000 | 0 | 0 | 0 | low pScore |
| 4 | | 2.591 | 2 | 0 | 0 | 0.000 | 0 | 0 | 0 | low pScore |
| 5 | | 1.501 | 3 | 0 | 0 | 0.370 | 1 | 0 | 0 | low pScore |
| 6 | | 0.370 | 1 | 0 | 0 | 0.370 | 1 | 0 | 0 | low pScore |
| 7 | | 1.796 | 4 | 0 | 0 | 0.000 | 0 | 0 | 0 | low pScore |
| 8 | | 1.427 | 2 | 0 | 0 | 0.000 | 0 | 0 | 0 | low pScore |
| 9 | | 1.587 | 4 | 0 | 0 | 0.457 | 2 | 0 | 0 | low pScore |
| 10 | | 4.397 | 2 | 0 | 0 | 2.834 | 1 | 0 | 0 | low pScore |
| 11 | | 0.000 | 0 | 0 | 0 | 0.000 | 0 | 0 | 0 | low pScore |
| 12 | | 2.787 | 2 | 0 | 0 | 0.367 | 1 | 0 | 0 | low pScore |
| 13 | | 0.095 | 1 | 0 | 0 | 0.095 | 1 | 0 | 0 | low pScore |
| 14 | | 1.562 | 1 | 0 | 0 | 0.000 | 0 | 0 | 0 | low pScore |
| 15 | | 4.378 | 2 | 0 | 0 | 0.000 | 0 | 0 | 0 | low pScore |
| 16 | | 2.860 | 4 | 0 | 0 | 0.000 | 0 | 0 | 0 | low pScore |
| 17 | | 0.000 | 0 | 0 | 0 | 0.000 | 0 | 0 | 0 | low pScore |

ii. **SmartsFilter** - This tool will help provide filtering based on reactivity. Simply paste one or more **compound node** SMILES directly from the CARLSBAD plugin into the field on the left of the screen, select at least one filter from the middle window pane (as you can see "Oprea" is the only filter selected in the screenshot below) and select the button "run smartsfilter." You will then be provided with information on reactivity in the "RESULTS" at the bottom of the page.

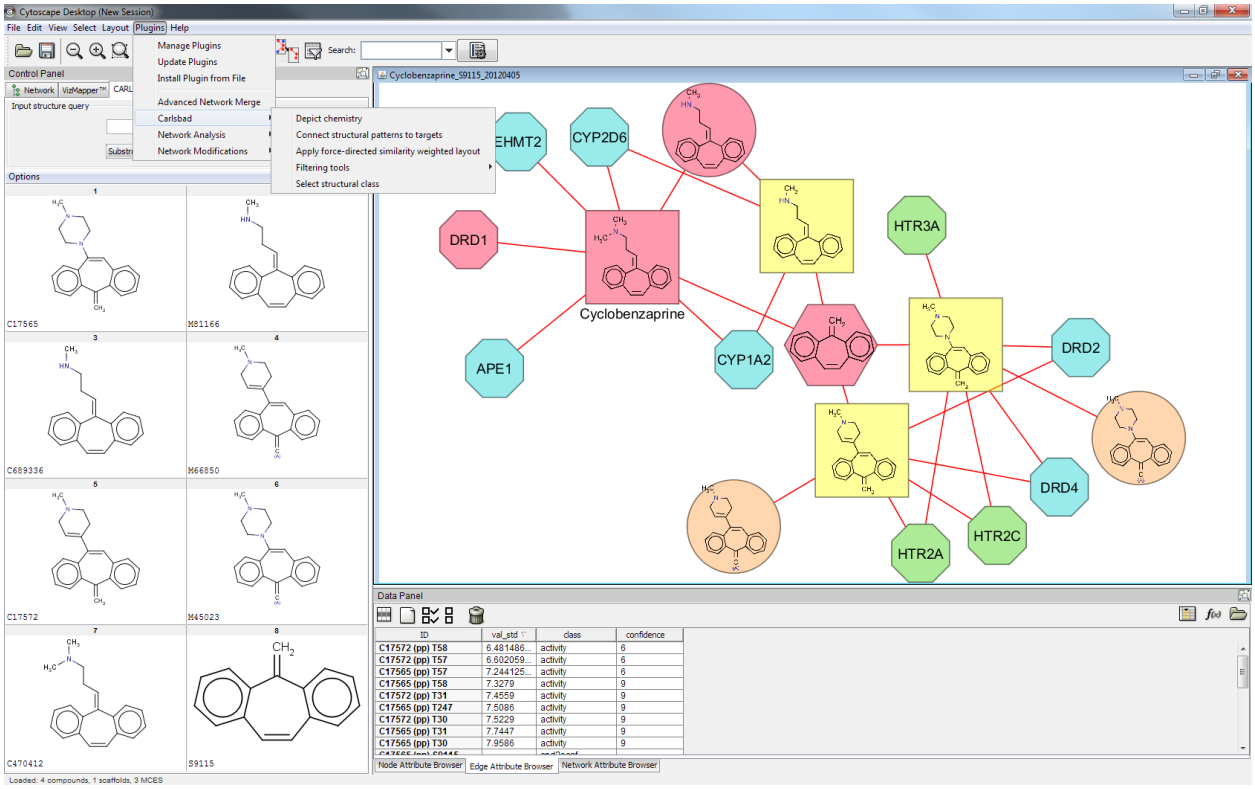



More information on the "PAINS" and "ALARM NMR" filters can be found in the following publications:

- Baell, J.B. & Holloway, G.A. New Substructure Filters for Removal of Pan Assay Interference Compounds (PAINS) from Screening Libraries and for Their Exclusion in Bioassays. *Journal of Medicinal Chemistry* **53**, 2719-2740 (2010).
- Huth, J.R. et al. ALARM NMR: A rapid and robust experimental method to detect reactive false positives in biochemical screens. *Journal of the American Chemical Society* **127**, 217-224 (2005).

## 6. Validation:

As a means of validating CARLSBAD, we looked at recent findings linking cyclobenzaprine to serotonin syndrome (SS) (Mestres *et al*., 2011). In the paper, Mestres showed that cyclobenzaprine probably bound several serotonin receptors causing the syndrome. We validated these findings with CARLSBAD by first performing a search for cyclobenzaprine in SNAKE, which revealed an association in the CARLSBAD database with only one target, DRD1, a dopamine receptor. After importing the XGMML file into Cytoscape via the CARLSBAD plugins automated importing method, *Load CARLSBAD network from SNAKE*, the network revealed two CCPs associated with cyclobenzaprine. Using the *Merge CARLSBAD networks* function to merge the cyclobenzaprine network with the networks of these CCPs revealed nine additional targets including two additional dopamine receptors, two cytochrome P450s, two enzymes and three serotonin receptors which were mentioned as likely being the causative agents of SS in Mestres's work. Thus, we were quickly able to predict off-target side effects that matched literature using CARLSBAD.



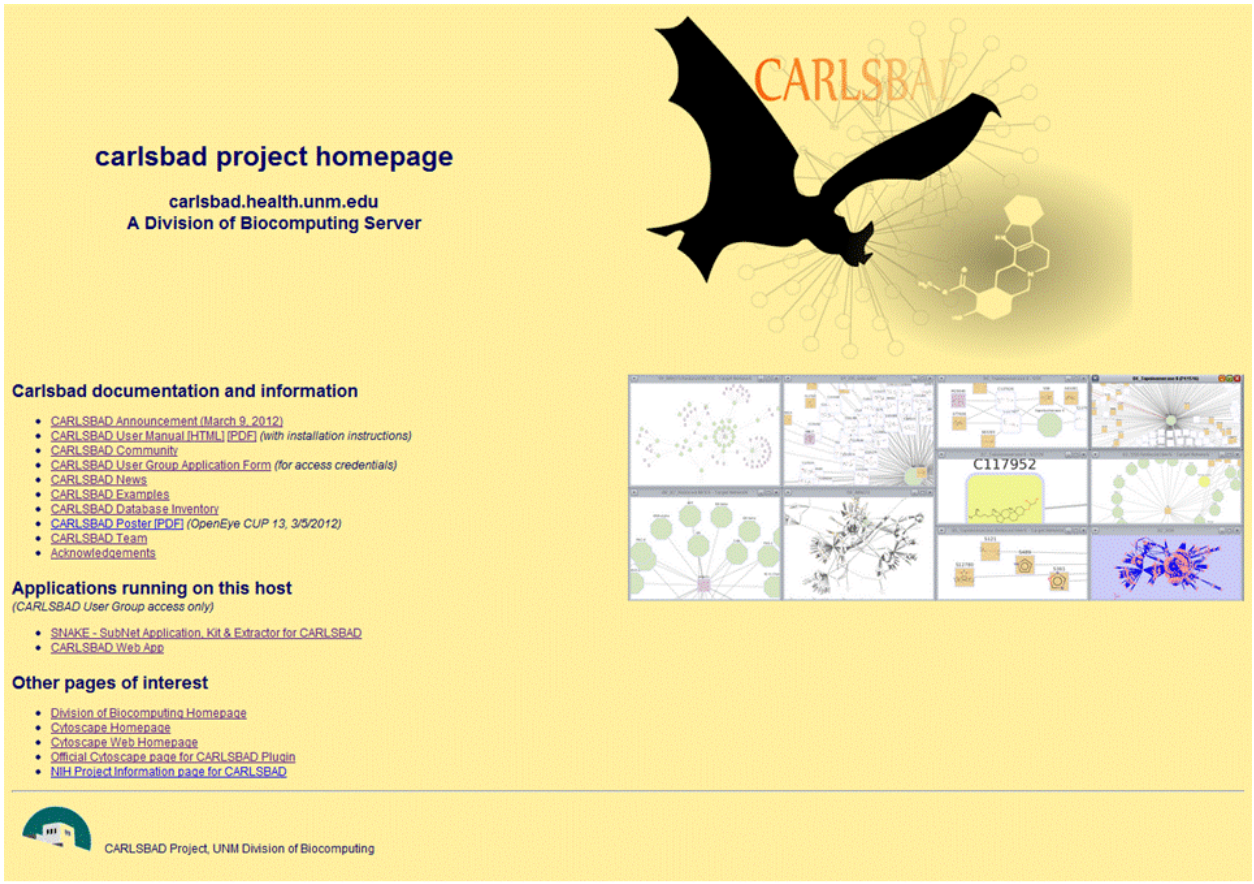


- Mestres, J., Seifert, S.A. and Oprea, T.I. (2011) Linking Pharmacology to Clinical Reports: Cyclobenzaprine and Its Possible Association With Serotonin Syndrome, *Clinical Pharmacology & Therapeutics*, **90**, 662-665.

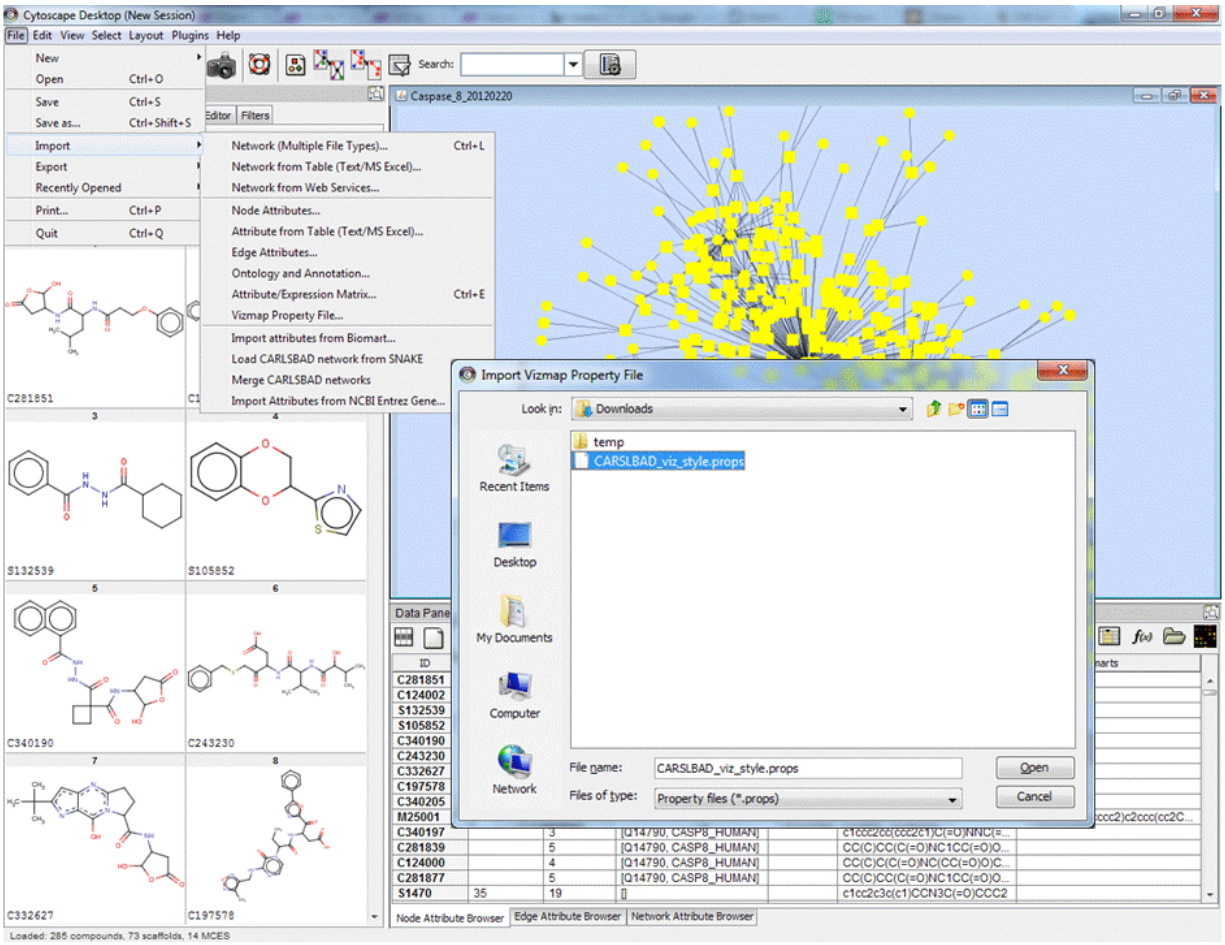
## 7. Appendix:

11. **Applying Different Visual Styles to a CARLSBAD Network** - It is possible to alter the visual style of your network by changing the settings in the “VizMapper” tab in the “Control Panel.” However, to facilitate this process we have designed “Vizmap Property Files” that you can use to enhance your network view.

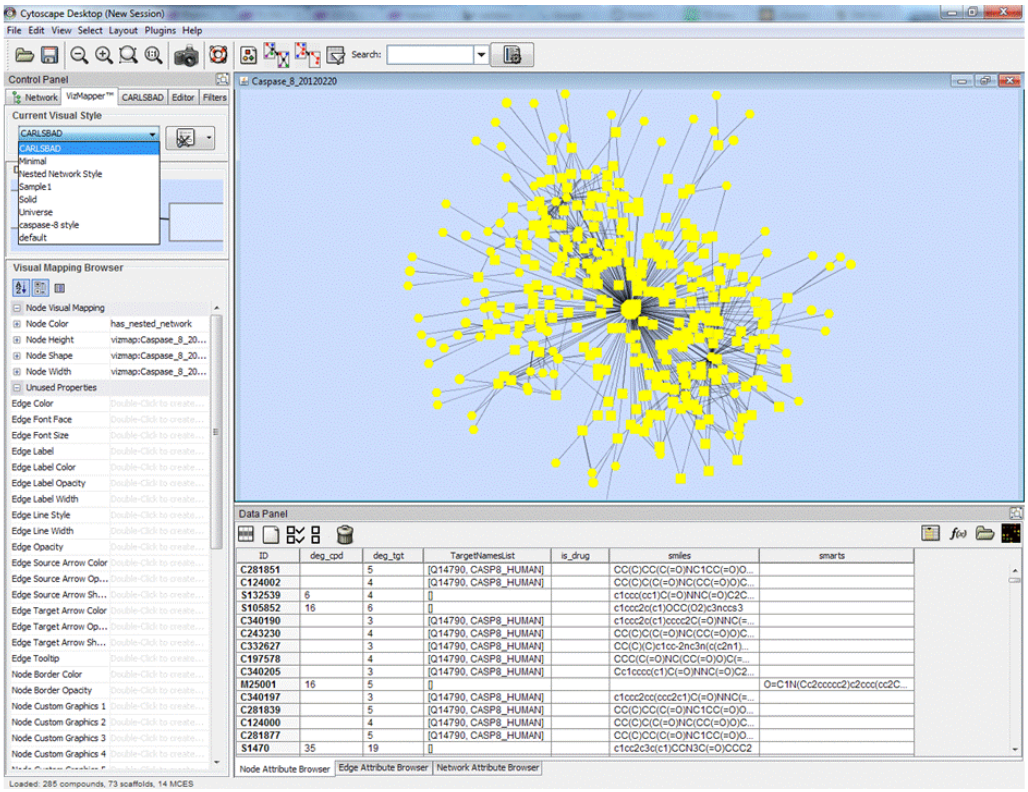
    - To accomplish this, simply download the “Vizmap Property File” from the CARLSBAD website **(carlsbad.health.unm.edu).** This file can be found in the "Downloads" section of the "CARLSBAD Community" link on the homepage.


carlsbad project homepage
carlsbad.health.unm.edu
A Division of Biocomputing Server
CARLSBA
C117952
Carlsbad documentation and information
CARLSBAD Announcement (March 9, 2012)
CARLSBAD User Manual [HTML] [PDF] (with installation instructions)
CARLSBAD Community
CARLSBAD User Group Application Form (for access credentials)
CARLSBAD News
CARLSBAD Examples
CARLSBAD Database Inventory
CARLSBAD Poster [PDF] (OpenEye CUP 13, 3/5/2012)
CARLSBAD Team
Acknowledgements
Applications running on this host
(CARLSBAD User Group access only)
SNAKE - SubNet Application, Kit & Extractor for CARLSBAD
CARLSBAD Web App
Other pages of interest
Division of Biocomputing Homepage
Cytoscape Homepage
Cytoscape Web Homepage
Official Cytoscape page for CARLSBAD Plugin
NIH Project Information page for CARLSBAD
CARLSBAD Project, UNM Division of Biocomputing

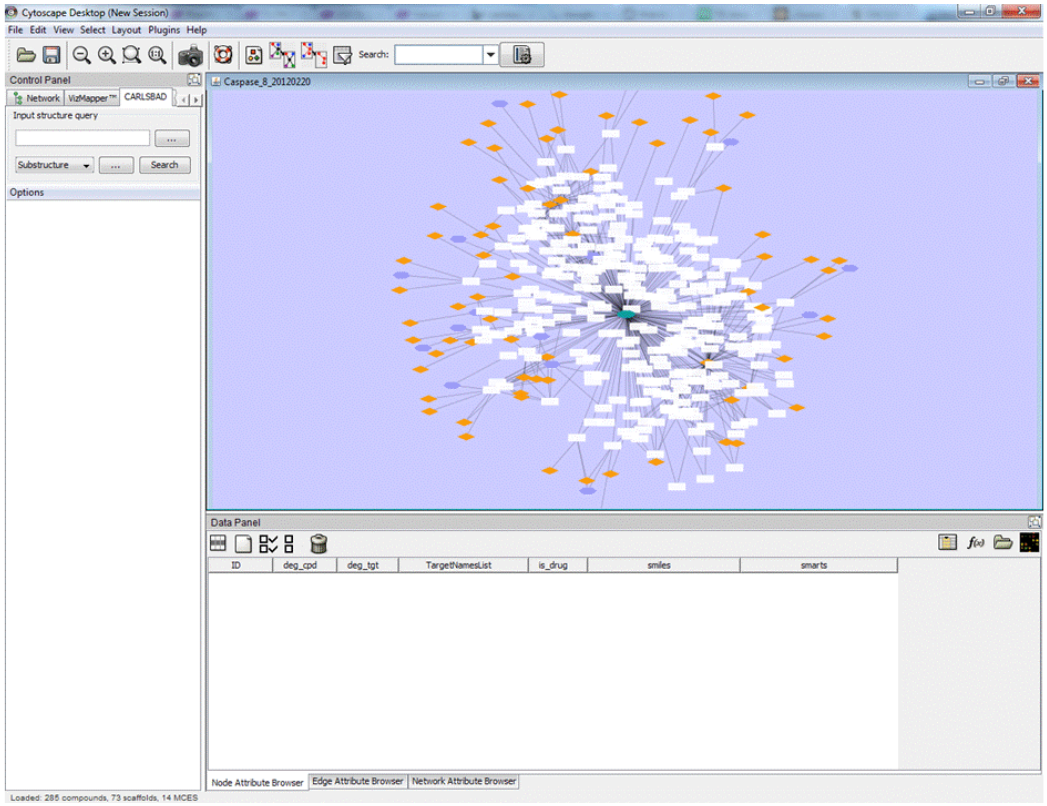
- After downloading the file, go to the “File” menu in Cytoscape and select “Import” and “Vizmap Property File.” Choose the file you downloaded from the CARLSBAD website in the pop-up window and select the “Open” button …




- …. Next go to the “VizMapper” tab to the left of the “CARLSBAD” tab in the “Control Panel.” You will see that the “CARLSBAD” visual style is selected, however to apply the new style that you downloaded, you must first choose another style, i.e. “Minimal,” and then select the CARLSBAD Visual style again ….




- … And the new VizMapper style you downloaded will be applied.



## 8. Glossary:

- **Apply force-directed similarity weighted layout** - A CARLSBAD plugin tool that contains options allowing the user to rearrange their network so that chosen node classes are grouped based on structural similarity and edge distance between the target(s) and compound(s) is related to the affinity the compound(s) have for the target(s). This is accomplished by changing the distance of compound edges connected to targets and CCPs based upon Tanimoto similarity values 0..1. In addition, Compound-Target edges, which have "val_std" activities are mapped to Tanimoto values: val_std >=8 -> 1; val_std < 8 and val_std > 5 -> val_std/10; val_std <= 5 -> 0.1.

- **Attribute** - A node/edge property.

- **CARLSBAD Plugin -** A Cytoscape plugin developed specifically for use with the CARLSBAD system.

- **CCP (Common Chemical Pattern)** - A term used to group MCESs and Scaffolds together as they are different means to the same end, common chemical patterns.

- **Compound (CPD)** - A small molecule or peptide that has the potential to modulate a target.

- **Connect structural patterns to targets** - A tool that connects the common chemical patterns (CCP) a compound is associated with directly to the target(s) by removing the compound nodes and creating a new edge directly to the target(s).

- **Control Panel** - A panel towards the left of the screen in Cytoscape that depicts structures in the CARLSBAD plugin and provides further Cytoscape options such as network visualization tools and custom filters.

- **Current Network** - The network you are actively investigating. To be contrasted to a network you might use to grow your current network.

- **Cytoscape** - Open source, community based, general purpose network analysis application, with many domain-specific plugins available.

- **Cytoscape Web** - A Flash plugin developed by the Cytoscape community for in-browser network analysis.

- **Data Panel** - A panel towards the bottom of the screen in Cytoscape that depicts attribute information and allows sorting by attribute.

- **deg_cpd** - The number of compounds another node (i.e., scaffold, MCES or target) is associated with in the CARLSBAD database, which may be in contrast to the number of compounds another node may be associated with in the current network.

- **deg_tgt** - The number of targets another node (i.e., Scaffold, MCES or Compound) is associated with in the CARLSBAD database, which may be in contrast to the number of targets another node may be associated with in the current network.

- **Depict Chemistry** - A way to view the chemical structure of a compound or common chemical pattern (CCP) in the selected nodes.

- **Disconnected Node** - A node that is not connected via an edge to any other nodes.

- **Edges** - Connections/relations between two nodes, represented by lines in CARLSBAD.

- **Edge Attribute Browser** - A tab at the bottom of the "Data Panel" specific to edge attributes.

- **Identifier** - A common ID that identifies a biological target such as a protein or gene; i.e., an EC number, ChEMBL ID, an Entrez Gene ID, an NCBI GI number, a RefSeq ID, a Uniprot ID, etc.

- **Lead Compound** - A compound that has a biological activity for a target and can be used as a starting point for chemical modifications that may lead to drug development.

- **MCES (Maximum Common Edge Substructure)** - An algorithm for identifying common chemical patterns. Sometimes referred to along with a scaffold as a CCP.

- **Network** - Any set of nodes and edges.

- **Network growth or expansion** - A term used to describe what happens to an initial network when another network is merged with it.

- **Network View Panel** - A panel towards the upper right of the screen in Cytoscape that depicts the network.

- **Node Attribute Browser** - A tab at the bottom of the "Data Panel" specific to node attributes.

- **Nodes** - A set of objects in a network (i.e., Compounds, Scaffolds, MCES and Targets), represented by shapes in CARLSBAD. A node may have zero or any number of associated edges.

- **Remove promiscuous compounds** - A filter that removes compounds above a specified "deg_tgt."

- **Remove specified activity edges** - A filter that removes edges below a defined "val_std" activity.

- **Remove unmergeable CCPs** - A filter that removes common chemical patterns (CCP) where the edges connected to the CCP in the network are equal to the "deg_cpd" of that CCP. In other words, a filter that removes CCPs with no potential for "network growth."

- **Repurposed Drug** - A new indication for a known drug.

- **Scaffold** - A means of identifying common chemical patterns via a collection of ring/ring systems in a compound and linkers which connect them. Sometimes referred to along with an MCES as a CCP.

- **Select structural class** - A tool that allows you to select the specific structural class of nodes in a network; i.e., compounds, scaffolds or MCESs.

- **SMARTS** - Line notation for specifying chemical substructure patterns.

- **SMILES** - Line notation for specifying chemical structures.

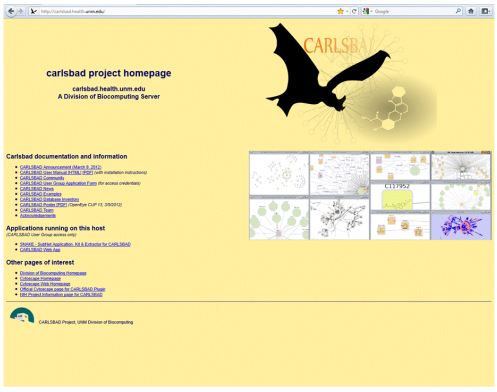
- **SNAKE (SubNet Application, Kit and Extractor)** - A web app for selecting a subset (sub-network) of the global CARLSBAD interaction network, and either (1) exporting that sub-net (XGMML) for import by Cytoscape, or (2) viewing and analyzing via Cytoscape Web.

- **Sub-Network** - Any subset of a larger network.

- **Target** - A biological target, most likely associated with some pathology, such as a protein involved in the apoptotic pathway.

- **Target of Interest (TOI)** - The target that is the focus of a network analysis.

- **val_std** - A bioactivity that has been "- log molar" transformed so that a "val_std" activity < 5.0 is the same as a an IC50 or EC50 value > 10 uM.



## 9. CARLSBAD References:


1. Bolton, E.E., *et al.* (2008) PubChem: Integrated Platform of Small Molecules and Biological Activities, *Annual Reports in Computational Chemistry*, **4**, 217-241.
2. CHORD, version 1.3, gNova, Inc., San Diego, CA, USA, www.gnova.com/, 2009.
3. Gaulton, A., *et al.* (2012) ChEMBL: a large-scale bioactivity database for drug discovery, *Nucleic Acids Research*, **40**, D1100-D1107.
4. Jaccard, P. (1901) Étude comparative de la distribution florale dans une portion des Alpes et des Jura, *Bulletin de la Société Vaudoise des Sciences Naturelles,* **37**, 547–579.
5. JChem Base, version 5.7, was used for cheminformatics processing, JChem 5.5, 2011, ChemAxon (http://www.chemaxon.com).
6. Lopes, C.T., *et al.* (2010) Cytoscape Web: an interactive web-based network browser, *Bioinformatics*, **26**, 2347-2348.
7. OEChem, version 1.7.4, OpenEye Scientific Software, Inc., Santa Fe, NM, USA, www.eyesopen.com, 2010.
8. Olah, M., *et al.* (2008) WOMBAT and WOMBAT-PK: Bioactivity Databases for Lead and Drug Discovery, In *Chemical Biology: From Small Molecules to Systems Biology and Drug Design*, Schreiber, S.L., Kapoor, T.M. and Wess G. (Eds), Wiley-VCH Verlag GmbH, Weinheim, Germany, pp 760-786.
9. Raymond, J.W., Gardiner, E.J. and Willett, P. (2002) RASCAL: Calculation of graph similarity using maximum common edge subgraphs, *Computer Journal*, **45**, 631-644.
10. Roth, B.L., *et al.* (2000) The Multiplicity of Serotonin Receptors: Uselessly Diverse Molecules or an Embarrassment of Riches?, *Neuroscientist*, **6**, 252-262.
11. Shannon, P., *et al.* (2003) Cytoscape: A software environment for integrated models of biomolecular interaction networks, *Genome Research*, **13**, 2498-2504.
12. Sharman, J.L., *et al.* (2011) IUPHAR-DB: new receptors and tools for easy searching and visualization of pharmacological data, *Nucl. Acids Res.*, **39**, D534-D538.
13. Yang, J.J. (2011) UNM-Biocomp-hscaf, Java library for HierS chemical scaffolds. Google Code open source project, http://code.google.com/p/unm-biocomp-hscaf.
14. Wilkens, S.J., Janes, J.J. and Su, A.I. (2004) Hiers: Hierarchical scaffold clustering, *Abstracts of Papers of the American Chemical Society*, **228**, U368-U368.

## 10. GNU Free Documentation License:

Version 1.3, 3 November 2008

Copyright © 2000, 2001, 2002, 2007, 2008 Free Software Foundation, Inc. <http://fsf.org/>

Everyone is permitted to copy and distribute verbatim copies of this license document, but changing it is not allowed.

***0. PREAMBLE***

The purpose of this License is to make a manual, textbook, or other functional and useful document "free" in the sense of freedom: to assure everyone the effective freedom to copy and redistribute it, with or without modifying it, either commercially or noncommercially. Secondarily, this License preserves for the author and publisher a way to get credit for their work, while not being considered responsible for modifications made by others.

This License is a kind of "copyleft", which means that derivative works of the document must themselves be free in the same sense. It complements the GNU General Public License, which is a copyleft license designed for free software.

We have designed this License in order to use it for manuals for free software, because free software needs free documentation: a free program should come with manuals providing the same freedoms that the software does. But this License is not limited to software manuals; it can be used for any textual work, regardless of subject matter or whether it is published as a printed book. We recommend this License principally for works whose purpose is instruction or reference.

***1. APPLICABILITY AND DEFINITIONS***

This License applies to any manual or other work, in any medium, that contains a notice placed by the copyright holder saying it can be distributed under the terms of this License. Such a notice grants a world-wide, royalty-free license, unlimited in duration, to use that work under the conditions stated herein. The "Document", below, refers to any such manual or work. Any member of the public is a licensee, and is addressed as "you". You accept the license if you copy, modify or distribute the work in a way requiring permission under copyright law.

A "Modified Version" of the Document means any work containing the Document or a portion of it, either copied verbatim, or with modifications and/or translated into another language.

A "Secondary Section" is a named appendix or a front-matter section of the Document that deals exclusively with the relationship of the publishers or authors of the Document to the Document's overall subject (or to related matters) and contains nothing that could fall directly within that overall subject. (Thus, if the Document is in part a textbook of mathematics, a Secondary Section may not explain any mathematics.) The relationship could be a matter of historical connection with the subject or with related matters, or of legal, commercial, philosophical, ethical or political position regarding them.

The "Invariant Sections" are certain Secondary Sections whose titles are designated, as being those of Invariant Sections, in the notice that says that the Document is released under this License. If a section does not fit the above definition of Secondary then it is not allowed to be designated as Invariant. The Document may contain zero Invariant Sections. If the Document does not identify any Invariant Sections then there are none.

The "Cover Texts" are certain short passages of text that are listed, as Front-Cover Texts or Back-Cover Texts, in the notice that says that the Document is released under this License. A Front-Cover Text may be at most 5 words, and a Back-Cover Text may be at most 25 words.

A "Transparent" copy of the Document means a machine-readable copy, represented in a format whose specification is available to the general public, that is suitable for revising the document straightforwardly with generic text editors or (for images composed of pixels) generic paint programs or (for drawings) some widely available drawing editor, and that is suitable for input to text formatters or for automatic translation to a variety of formats suitable for input to text formatters. A copy made in an otherwise Transparent file format whose markup, or absence of markup, has been arranged to thwart or discourage subsequent modification by readers is not Transparent. An image format is not Transparent if used for any substantial amount of text. A copy that is not "Transparent" is called "Opaque".

Examples of suitable formats for Transparent copies include plain ASCII without markup, Texinfo input format, LaTeX input format, SGML or XML using a publicly available DTD, and standard-conforming simple HTML, PostScript or PDF designed for human modification. Examples of transparent image formats include PNG, XCF and JPG. Opaque formats include proprietary formats that can be read and edited only by proprietary word processors, SGML or XML for which the DTD and/or processing tools are not generally available, and the machine-generated HTML, PostScript or PDF produced by some word processors for output purposes only.

The "Title Page" means, for a printed book, the title page itself, plus such following pages as are needed to hold, legibly, the material this License requires to appear in the title page. For works in formats which do not have any title page as such, "Title Page" means the text near the most prominent appearance of the work's title, preceding the beginning of the body of the text.

The "publisher" means any person or entity that distributes copies of the Document to the public.

A section "Entitled XYZ" means a named subunit of the Document whose title either is precisely XYZ or contains XYZ in parentheses following text that translates XYZ in another language. (Here XYZ stands for a specific section name mentioned below, such as "Acknowledgements", "Dedications", "Endorsements", or "History".) To "Preserve the Title" of such a section when you modify the Document means that it remains a section "Entitled XYZ" according to this definition.

The Document may include Warranty Disclaimers next to the notice which states that this License applies to the Document. These Warranty Disclaimers are considered to be included by reference in this License, but only as regards disclaiming warranties: any other implication that these Warranty Disclaimers may have is void and has no effect on the meaning of this License.

***2. VERBATIM COPYING***

You may copy and distribute the Document in any medium, either commercially or noncommercially, provided that this License, the copyright notices, and the license notice saying this License applies to the Document are reproduced in all copies, and that you add no other conditions whatsoever to those of this License. You may not use technical measures to obstruct or control the reading or further copying of the copies you make or distribute. However, you may accept compensation in exchange for copies. If you distribute a large enough number of copies you must also follow the conditions in section 3.

You may also lend copies, under the same conditions stated above, and you may publicly display copies.

***3. COPYING IN QUANTITY***

If you publish printed copies (or copies in media that commonly have printed covers) of the Document, numbering more than 100, and the Document's license notice requires Cover Texts, you must enclose the copies in covers that carry, clearly and legibly, all these Cover Texts: Front-Cover Texts on the front cover, and Back-Cover Texts on the back cover. Both covers must also clearly and legibly identify you as the publisher of these copies. The front cover must present the full title with all words of the title equally prominent and visible. You may add other material on the covers in addition. Copying with changes limited to the covers, as long as they preserve the title of the Document and satisfy these conditions, can be treated as verbatim copying in other respects.

If the required texts for either cover are too voluminous to fit legibly, you should put the first ones listed (as many as fit reasonably) on the actual cover, and continue the rest onto adjacent pages.

If you publish or distribute Opaque copies of the Document numbering more than 100, you must either include a machine-readable Transparent copy along with each Opaque copy, or state in or with each Opaque copy a computer-network location from which the general network-using public has access to download using public-standard network protocols a complete Transparent copy of the Document, free of added material. If you use the latter option, you must take reasonably prudent steps, when you begin distribution of Opaque copies in quantity, to ensure that this Transparent copy will remain thus accessible at the stated location until at least one year after the last time you distribute an Opaque copy (directly or through your agents or retailers) of that edition to the public.

It is requested, but not required, that you contact the authors of the Document well before redistributing any large number of copies, to give them a chance to provide you with an updated version of the Document.

***4. MODIFICATIONS***

You may copy and distribute a Modified Version of the Document under the conditions of sections 2 and 3 above, provided that you release the Modified Version under precisely this License, with the Modified Version filling the role of the Document, thus licensing distribution and modification of the Modified Version to whoever possesses a copy of it. In addition, you must do these things in the Modified Version:

- A. Use in the Title Page (and on the covers, if any) a title distinct from that of the Document, and from those of previous versions (which should, if there were any, be listed in the History section of the Document). You may use the same title as a previous version if the original publisher of that version gives permission.
- B. List on the Title Page, as authors, one or more persons or entities responsible for authorship of the modifications in the Modified Version, together with at least five of the principal authors of the Document (all of its principal authors, if it has fewer than five), unless they release you from this requirement.
- C. State on the Title page the name of the publisher of the Modified Version, as the publisher.
- D. Preserve all the copyright notices of the Document.
- E. Add an appropriate copyright notice for your modifications adjacent to the other copyright notices.
- F. Include, immediately after the copyright notices, a license notice giving the public permission to use the Modified Version under the terms of this License, in the form shown in the Addendum below.
- G. Preserve in that license notice the full lists of Invariant Sections and required Cover Texts given in the Document's license notice.
- H. Include an unaltered copy of this License.
- I. Preserve the section Entitled "History", Preserve its Title, and add to it an item stating at least the title, year, new authors, and publisher of the Modified Version as given on the Title Page. If there is no section Entitled "History" in the Document, create one stating the title, year, authors, and publisher of the Document as given on its Title Page, then add an item describing the Modified Version as stated in the previous sentence.
- J. Preserve the network location, if any, given in the Document for public access to a Transparent copy of the Document, and likewise the network locations given in the Document for previous versions it was based on. These may be placed in the "History" section. You

may omit a network location for a work that was published at least four years before the Document itself, or if the original publisher of the version it refers to gives permission.

- K. For any section Entitled "Acknowledgements" or "Dedications", Preserve the Title of the section, and preserve in the section all the substance and tone of each of the contributor acknowledgements and/or dedications given therein.
- L. Preserve all the Invariant Sections of the Document, unaltered in their text and in their titles. Section numbers or the equivalent are not considered part of the section titles.
- M. Delete any section Entitled "Endorsements". Such a section may not be included in the Modified Version.
- N. Do not retitle any existing section to be Entitled "Endorsements" or to conflict in title with any Invariant Section.
- O. Preserve any Warranty Disclaimers.

If the Modified Version includes new front-matter sections or appendices that qualify as Secondary Sections and contain no material copied from the Document, you may at your option designate some or all of these sections as invariant. To do this, add their titles to the list of Invariant Sections in the Modified Version's license notice. These titles must be distinct from any other section titles.

You may add a section Entitled "Endorsements", provided it contains nothing but endorsements of your Modified Version by various parties—for example, statements of peer review or that the text has been approved by an organization as the authoritative definition of a standard.

You may add a passage of up to five words as a Front-Cover Text, and a passage of up to 25 words as a Back-Cover Text, to the end of the list of Cover Texts in the Modified Version. Only one passage of Front-Cover Text and one of Back-Cover Text may be added by (or through arrangements made by) any one entity. If the Document already includes a cover text for the same cover, previously added by you or by arrangement made by the same entity you are acting on behalf of, you may not add another; but you may replace the old one, on explicit permission from the previous publisher that added the old one.

The author(s) and publisher(s) of the Document do not by this License give permission to use their names for publicity for or to assert or imply endorsement of any Modified Version.

***5. COMBINING DOCUMENTS***

You may combine the Document with other documents released under this License, under the terms defined in section 4 above for modified versions, provided that you include in the combination all of the Invariant Sections of all of the original documents, unmodified, and list them all as Invariant Sections of your combined work in its license notice, and that you preserve all their Warranty Disclaimers.

The combined work need only contain one copy of this License, and multiple identical Invariant Sections may be replaced with a single copy. If there are multiple Invariant Sections with the same name but different contents, make the title of each such section unique by adding at the end of it, in parentheses, the name of the original author or publisher of that section if known, or else a unique number. Make the same adjustment to the section titles in the list of Invariant Sections in the license notice of the combined work.

In the combination, you must combine any sections Entitled "History" in the various original documents, forming one section Entitled "History"; likewise combine any sections Entitled "Acknowledgements", and any sections Entitled "Dedications". You must delete all sections Entitled "Endorsements".

***6. COLLECTIONS OF DOCUMENTS***

You may make a collection consisting of the Document and other documents released under this License, and replace the individual copies of this License in the various documents with a single copy that is included in the collection, provided that you follow the rules of this License for verbatim copying of each of the documents in all other respects.

You may extract a single document from such a collection, and distribute it individually under this License, provided you insert a copy of this License into the extracted document, and follow this License in all other respects regarding verbatim copying of that document.

***7. AGGREGATION WITH INDEPENDENT WORKS***

A compilation of the Document or its derivatives with other separate and independent documents or works, in or on a volume of a storage or distribution medium, is called an "aggregate" if the copyright resulting from the compilation is not used to limit the legal rights of the compilation's users beyond what the individual works permit. When the Document is included in an aggregate, this License does not apply to the other works in the aggregate which are not themselves derivative works of the Document.

If the Cover Text requirement of section 3 is applicable to these copies of the Document, then if the Document is less than one half of the entire aggregate, the Document's Cover Texts may be placed on covers that bracket the Document within the aggregate, or the electronic equivalent of covers if the Document is in electronic form. Otherwise they must appear on printed covers that bracket the whole aggregate.

***8. TRANSLATION***

Translation is considered a kind of modification, so you may distribute translations of the Document under the terms of section 4. Replacing Invariant Sections with translations requires special permission from their copyright holders, but you may include translations of some or all Invariant Sections in addition to the original versions of these Invariant Sections. You may include a translation of this License, and all the license notices in the Document, and any Warranty Disclaimers, provided that you also include the original English version of this License and

the original versions of those notices and disclaimers. In case of a disagreement between the translation and the original version of this License or a notice or disclaimer, the original version will prevail.

If a section in the Document is Entitled "Acknowledgements", "Dedications", or "History", the requirement (section 4) to Preserve its Title (section 1) will typically require changing the actual title.

***9. TERMINATION***

You may not copy, modify, sublicense, or distribute the Document except as expressly provided under this License. Any attempt otherwise to copy, modify, sublicense, or distribute it is void, and will automatically terminate your rights under this License.

However, if you cease all violation of this License, then your license from a particular copyright holder is reinstated (a) provisionally, unless and until the copyright holder explicitly and finally terminates your license, and (b) permanently, if the copyright holder fails to notify you of the violation by some reasonable means prior to 60 days after the cessation.

Moreover, your license from a particular copyright holder is reinstated permanently if the copyright holder notifies you of the violation by some reasonable means, this is the first time you have received notice of violation of this License (for any work) from that copyright holder, and you cure the violation prior to 30 days after your receipt of the notice.

Termination of your rights under this section does not terminate the licenses of parties who have received copies or rights from you under this License. If your rights have been terminated and not permanently reinstated, receipt of a copy of some or all of the same material does not give you any rights to use it.

***10. FUTURE REVISIONS OF THIS LICENSE***

The Free Software Foundation may publish new, revised versions of the GNU Free Documentation License from time to time. Such new versions will be similar in spirit to the present version, but may differ in detail to address new problems or concerns. See http://www.gnu.org/copyleft/.

Each version of the License is given a distinguishing version number. If the Document specifies that a particular numbered version of this License "or any later version" applies to it, you have the option of following the terms and conditions either of that specified version or of any later version that has been published (not as a draft) by the Free Software Foundation. If the Document does not specify a version number of this License, you may choose any version ever published (not as a draft) by the Free Software Foundation. If the Document specifies that a proxy can decide which future versions of this License can be used, that proxy's public statement of acceptance of a version permanently authorizes you to choose that version for the Document.

***11. RELICENSING***

"Massive Multiauthor Collaboration Site" (or "MMC Site") means any World Wide Web server that publishes copyrightable works and also provides prominent facilities for anybody to edit those works. A public wiki that anybody can edit is an example of such a server. A "Massive Multiauthor Collaboration" (or "MMC") contained in the site means any set of copyrightable works thus published on the MMC site.

"CC-BY-SA" means the Creative Commons Attribution-Share Alike 3.0 license published by Creative Commons Corporation, a not-for-profit corporation with a principal place of business in San Francisco, California, as well as future copyleft versions of that license published by that same organization.

"Incorporate" means to publish or republish a Document, in whole or in part, as part of another Document.

An MMC is "eligible for relicensing" if it is licensed under this License, and if all works that were first published under this License somewhere other than this MMC, and subsequently incorporated in whole or in part into the MMC, (1) had no cover texts or invariant sections, and (2) were thus incorporated prior to November 1, 2008.

The operator of an MMC Site may republish an MMC contained in the site under CC-BY-SA on the same site at any time before August 1, 2009, provided the MMC is eligible for relicensing.